

OY	481	TCGCTGCTCCAAAATTGCAATGGAAGGAGGCTCCCTAACACAGCATCGTTTCAAAGTT	540
Db	481	TCGCTGCTCCAAAATTGCAATGGAAGGAGGAGGCTCCCTAACACAGCATCGTTTCAAAGTT	540
OY	541	GGACACAGCTCTGAACATTTAAATTTATCTCCAAATCTACTGAATAACGGCTAGCGCGAAT	600
Db	541	GGACACAGCTCTGAACATTTAAATTTATCTCCAAATCTACTGAATAACGGCTAGCGCGAAT	600
OY	601	CAACAATAATGTGGACACGACGATGATGTATCTATCACAAAAGAAAGACTTTCCCGAATCG	660
Db	601	CAACAATAATGTGGACACGACGATGATGTATCTATCACAAAAGAAAGACTTTCCCGAATCG	660
OY	661	CTTCCAGTGTGAGATGACGACGACGAACAAGTCTTTAGCTGTGATGTCMAATATTTGGATAGAC	720
Db	661	CTTCCAGTGTGAGATGACGACGACGAACAAGTCTTTAGCTGTGATGTCMAATATTTGGATAGAC	720
OY	721	GAGGGAAGCTTCTTCATATTTCTGTGGATTTGCGAGCAGTCAATGATATCAACTAAAT	780
Db	721	GAGGGAAGCTTCTTCATATTTCTGTGGATTTGCGAGCAGTCAATGATATCAACTAAAT	780
OY	781	GGCAAAAAAGGAATATTTGAGCCACACAAATTTAAATGCTCAAGCAATTTAAAAATGGAGAA	840
Db	781	GGCAAAAAAGGAATATTTGAGCCACACAAATTTAAATGCTCAAGCAATTTAAAAATGGAGAA	840
OY	841	AATCCAGATATGTAGCTAGACGACGCTCACAATTTGATCAATTAGCAAAAAGATCCAAATGAT	900
Db	841	AATCCAGATATGTAGCTAGACGACGCTCACAATTTGATCAATTAGCAAAAAGATCCAAATGAT	900
OY	901	CCCATGATTTGGACAAAAGTTTAAGGTTGGACAAAAGTTTGAGCTATCGACCCCTTGAGCT	960
Db	901	CCCATGATTTGGACAAAAGTTTAAGGTTGGACAAAAGTTTGAGCTATCGACCCCTTGAGCT	960
OY	961	CAGCAATTCATTAACCTCCAGCTGCGCTCGCATTCGAAATTTGCGAAAACCTGAAGAGAT	1020
Db	961	CAGCAATTCATTAACCTCCAGCTGCGCTCGCATTCGAAATTTGCGAAAACCTGAAGAGAT	1020
OY	1021	CTTATTTGGGAATGATGATGCTCGAGATGACACTTGAAGACAGTTTCTTATTCATATCAAT	1080
Db	1021	CTTATTTGGGAATGATGATGCTCGAGATGACACTTGAAGACAGTTTCTTATTCATATCAAT	1080
OY	1081	AATACATTTATATGTTCCCAAGTTGTTATGCGGAAAAGTAAATTTTGGAACTTGTCGCCCA	1140
Db	1081	AATACATTTATATGTTCCCAAGTTGTTATGCGGAAAAGTAAATTTTGGAACTTGTCGCCCA	1140
OY	1141	GATAGATTTCAAGGAACTTACAGATGGGATGAATACTTGGAGAAAAGATTCGCAAGAAAC	1200
Db	1141	GATAGATTTCAAGGAACTTACAGATGGGATGAATACTTGGAGAAAAGATTCGCAAGAAAC	1200
OY	1201	CTACCGCTGACCTGTTCAACGCCAATGCTTCCCAAGAGATATAGACAAATTTTAAAGTTA	1260
Db	1201	CTACCGCTGACCTGTTCAACGCCAATGCTTCCCAAGAGATATAGACAAATTTTAAAGTTA	1260
OY	1261	ATTCTGATTTCCAAAAGGGTAGAGCTACGCTTGAGAGCTGCTGACATGTGTGAAAATCAG	1320
Db	1261	ATTCTGATTTCCAAAAGGGTAGAGCTACGCTTGAGAGCTGCTGACATGTGTGAAAATCAG	1320
OY	1321	TTTATTTGTCAGCTACAGTGAATCAAGTTATGATGGAAGACATGATTAATGTCAATTTTGAC	1380
Db	1321	TTTATTTGTCAGCTACAGTGAATCAAGTTATGATGGAAGACATGATTAATGTCAATTTTGAC	1380
OY	1381	GGCTGGGATGAAGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1440
Db	1381	GGCTGGGATGAAGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1440
OY	1441	GGATGGGTGAAGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1497
Db	1441	GGATGGGTGAAGGACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	1497

RESULT 3
US-09-872-523-74
; Sequence 74, Application US/09872523
; Patent No. US20020137906A1

```

: GENERAL INFORMATION:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Davison, Ewa M.
: APPLICANT: Lu, Xiaowei
: TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
: FILE REFERENCE: 01997/536002
: CURRENT APPLICATION NUMBER: US/09/872,523
: CURRENT FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: US 60/208,802
: PRIOR FILING DATE: 2000-06-02
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 74
: LENGTH: 1497
: TYPE: DNA
: ORGANISM: Caenorhabditis elegans
: US-09-872-523-74

```

Query Match	96.6%;	Score 1466;	DB 10;	Length 1497;
Local Similarity	99.9%;	Pred. No. 0;		
Matches 1496;	Conservative	0;	Mismatches	1;
			Indels	0;
			Gaps	0;
OY	1	ATGTCGAAATTTCTGAAAAATTTGTCAGAGCTAAACAAAAATCGAGAGAAACTCGATAAG	60	
Db	1	ATGTCGAAATTTCTGAAAAATTTGTCAGAGCTAAACAAAAATCGAGAGAAACTCGATAAG	60	
OY	61	ACCTAATTGTGGGAATCCTATTATACATAGTTCCGAAAGGAAAAACTTTCTTACTTCCA	120	
Db	61	ACCTAATTGTGGGAATCCTATTATACATAGTTCCGAAAGGAAAAACTTTCTTACTTCCA	120	
OY	121	GTTGAGACATTCACATCGTAACCTTACAGTTAAATTTAAAGAAATCGTAAGAGAGAGTT	180	
Db	121	GTTGAGACATTCACATCGTAACCTTACAGTTAAATTTAAAGAAATCGTAAGAGAGAGTT	180	
OY	181	ATCTTCGAACAGTGTGCTCATGATATTAGACAAGAACTCGCATTCGATTCAAGTCAGATGG	240	
Db	181	ATCTTCGAACAGTGTGCTCATGATATTAGACAAGAACTCGATTCGATTCAAGTCAGATGG	240	
OY	241	TTTGGAGCAATTTGAAAAAGTTTGGCGATACAGAGTTTGGGCTCATTTATGCGAGCTGAC	300	
Db	241	TTTGGAGCAATTTGAAAAAGTTTGGCGATACAGAGTTTGGGCTCATTTATGCGAGCTGAC	300	
OY	301	ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTTGGCAACGCCCCAAATG	360	
Db	301	ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTTGGCAACGCCCCAAATG	360	
OY	361	AGTATCCCAATATGAGATAAATGTATATGCTCCGCCGCTTGCAATCAACGAGAGATTC	420	
Db	361	AGTATCCCAATATGAGATAAATGTATATGCTCCGCCGCTTGCAATCAACGAGAGATTC	420	
OY	421	CAAAATGATATGATGTAATTTATGTAATTAATGCAATGATGAGCGAAATGTCGGCCAACT	480	
Db	421	CAAAATGATATGATGTAATTTATGTAATTAATGCAATGATGAGCGAAATGTCGGCCAACT	480	
OY	481	TCGCTGTCTCCAAAATTTGATGAGGAGGAGGCTCCTTAAGCAACATCGTTTCAAAGTT	540	
Db	481	TCGCTGTCTCCAAAATTTGATGAGGAGGAGGCTCCTTAAGCAACATCGTTTCAAAGTT	540	
OY	541	GGACACAGCTCTTGAACTATTAAATTTATTCCAATTTACTGAAPATACGGGTACGCGGAATT	600	
Db	541	GGACACAGCTCTTGAACTATTAAATTTATTCCAATTTACTGAAPATACGGGTACGCGGAATT	600	
OY	601	CAAGAAATATGTGGGACGACGAGATGAATATCTATACAAAAGAAAGACTTTCGGAAATCG	660	
Db	601	CAAGAAATATGTGGGACGACGAGATGAATATCTATACAAAAGAAAGACTTTCGGAAATCG	660	
OY	661	CTTCAGATGAGATGACGACAGCAAGCTTTTACCTCGATCTCAATATTTGGATAGAC	720	
Db	661	CTTCAGATGAGATGACGACAGCAAGCTTTTACCTCGATCTCAATATTTGGATAGAC	720	
OY	721	GAGGAGACCTTCTTCATATTTCTGCTGGATTTGACGACATGATGATATCACTAAAT	780	
Db	721	GAGGAGACCTTCTTCATATTTCTGCTGGATTTGACGACATGATGATATCACTAAAT	780	

Db	1021	CTTATTGTGGGAATGANGTGTCCAGATGCACTTTGAAGCACTTTTCCATTCAATTCAT	1080
OY	1081	AATACATTATGTTCCAGTGTGTTATGTGGAAAAATATAATTTTGAACCTGTGTTCCGCCA	1140
Db	1081	AATACATTATGTTCCAGTGTGTTATGTGGAAAAATATAATTTTGAACCTGTGTTCCGCCA	1140
OY	1141	GATAGATTCAAGGAACAATTCAGATGATGGATCAATATCTTGGAGAAAGAAATCTGCAGAAACC	1200
Db	1141	GATAGATTCAAGGAACAATTCAGATGATGGATCAATATCTTGGAGAAAGAAATCTGCAGAAACC	1200
OY	1201	CTACCGCTGACTGTGTTCCAAAGCCATGCTCTCCCAAGAGAAATTTAGCAAAATTTAAGSTA	1260
Db	1201	CTACCGCTGACTGTGTTCCAAAGCCATGCTCTCCCAAGAGAAATTTAGCAAAATTTAAGSTA	1260
OY	1261	ATTCTGATTTCCAAACGGGTAGGACATACGCTTAAGCTGCTGCATGTGTGAAAATCAG	1320
Db	1261	ATTCTGATTTCCAAACGGGTAGGACATACGCTTAAGCTGCTGCATGTGTGAAAATCAG	1320
OY	1321	TTTATTGTGTCAGCTACAGTGAATTCAGTTATGSAAGACATGAATATGTCATTTTGAC	1380
Db	1321	TTTATTGTGTCAGCTACAGTGAATTCAGTTATGSAAGACATGAATATGTCATTTTGAC	1380
OY	1381	GGCTGGGATGAAGAATTTGATGAACCTGTATGATGTGGACCTCCATGATATTCTACCGATA	1440
Db	1381	GGCTGGGATGAAGAATTTGATGAACCTGTATGATGTGGACCTCCATGATATTCTACCGATA	1440
OY	1441	GCATGGCTGTGAAGCGCACAGTTATGTTCTACAAACCTCGGAAAAAGTACAACTATTGGA	1497
Db	1441	GCATGGCTGTGAAGCGCACAGTTATGTTCTACAAACCTCGGAAAAAGTACAACTATTGGA	1497

```

RESULT 5
US-09-872-523-76
Sequence 76, Application US/09872523
Patent No. US20020137906A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Davison, Eva M.
APPLICANT: Lu, Xiaowei
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
FILE REFERENCE: 01997/536002
CURRENT APPLICATION NUMBER: US/09/872.523
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/208,802
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 2307
TYPE: DNA
ORGANISM: Caenorhabditis elegans
US-09-872-523-76

```

```

Query Match      55.6%; Score 833; DB 10; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	448	AATTTGCAATGATGGCGAAATCGTGGGCAAACTTCGCTGTCACAAATTCGATGAAGG	507
Db	1138	AATTTGCAATGATGGCGAAATGTTGGGCCAACTTCGCTGTCACAAATTCGATGAAGG	1197
QY	508	AAGGCTCTCCCAACCAAGACATCGTTTCAAAATTTGAGCAACGCTTTGAACATTAATAATAT	567
Db	1198	AAGGCTCTCCCAACCAAGACATCGTTTCAAAATTTGAGCAACGCTTTGAACATTAATAATAT	1257
QY	568	TCCAAATTCYACTGAAATACGGCGTAGCGCGAATTCAGAATAATATGTGACGACGATGAAT	627
Db	1258	TCCAAATTCYACTGAAATACGGCGTAGCGCGAATTCAGAATAATATGTGACGACGATGAAT	1317
QY	628	GTATCTATCACAAAGAAAGACTTCCCGGAATCGCTTCAGATGCAGATGACGACAGACAA	687
Db	1318	GTATCTATCACAAAGAAAGACTTTCGCCGAATCGCTTCAGATGCAGATGACGACAGACAA	1377

Oy	688	GTCTTTAGCCTCGATCTCATATTGGATTACACGAGAAACCTCTTCATATTCCGCTT	747
Db	1378	GTCTTTACTCTCGATCTCATATTGGATTACACGAGGAACCTCTTCATATTCCGCTT	1437
Oy	748	GGATTTTGACGAGTCATAGATGATTCACCTAAATGCGAAAAAGGATATTATTGACACACA	807
Db	1438	GGATTTTGACGAGTCATAGATGATTCACCTAAATGCGAAAAAGGATATTATTGACACACA	1497
Oy	808	AATAAATTTGCTCAAGCATATAAAAAATGAGAAAAATCCAGATATTGACTCAGACAGCTC	867
Db	1498	AATAAATTTGCTCAAGCATATAAAAAATGAGAAAAATCCAGATATTGACTCAGACAGCTC	1557
Oy	868	ACATTGTATTCAAATTAGCAAAAGATCCCATTTATCCCATGATTTGGAGAAAAAGTTAAGGTT	927
Db	1558	ACATTGTATTCAAATTAGCAAAAGATCCCATTTATCCCATGATTTGGAGAAAAAGTTAAGGTT	1617
Oy	928	GGACAAAAGTTTGAGCTCATCGACCCCTGGCTCAGCAATTCATTAACCTCCAGCTGCGCT	987
Db	1618	GGACAAAAGTTTGAGCTCATCGACCCCTGGCTCAGCAATTCATTAACCTCCAGCTGCGCT	1677
Oy	988	TCGATTTCTCAAAATTTTGCAAAAGTCGAAGGATATCTTATTGTGGGAATGAGATGTCAGAT	1047
Db	1678	TCGATTTCTCAAAATTTTGCAAAAGTCGAAGGATATCTTATTGTGGGAATGAGATGTCAGAT	1737
Oy	1048	GCACCTTGAAAGACATTTTCCATTCCATATCATCAATTAATACATTTAATGTTCCCGAGTGGTAT	1107
Db	1738	GCACCTTGAAAGACATTTTCCATTCCATATCATCAATTAATACATTTAATGTTCCCGAGTGGTAT	1797
Oy	1108	CGCGAAAAAGTTTAATTTTGGAACTGTTCCCGCAGATGAGTTCAAGCAACATTCAGATAGG	1167
Db	1798	CGCGAAAAAGTTTAATTTTGGAACTGTTCCCGCAGATGAGTTCAAGCAACATTCAGATAGG	1857
Oy	1168	GATGATATCTTGGAGAAAAGATTCGACAGAAACCTACCGCTTGACTTGTTCAGCCCAATG	1227
Db	1858	GATGATATCTTGGAGAAAAGATTCGACAGAAACCTACCGCTTGTTCAGCCCAATG	1917
Oy	1228	CCTTCCCAAGAGATTTAGACAATTTTAAGTATTTCGATTGCCAAACGGGT	1280
Db	1918	CCTTCCCAAGAGATTTAGACAATTTTAAGTATTTCGATTGCCAAACGGGT	1970

```

: RESULT 6
: US-09-872-523-77
: Sequence 77, Application US/09872523
: Patent No. US20020137906A1
: GENERAL INFORMATION:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Davison, Ewa M.
: APPLICANT: Lu, Xiaowei
: TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
: FILE REFERENCE: 01997/536002
: CURRENT APPLICATION NUMBER: US/09/872.523
: CURRENT FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: US 60/208,802
: PRIOR FILING DATE: 2000-06-02
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 77
: LENGTH: 2307
: TYPE: DNA
: ORGANISM: Caenorhabditis elegans
: US-09-872-523-77

```

	Query Match	Similarity	Score 833	DB 10	Length 2307
	Best Local	100.0%	Pred. No. 0		
	Matches 833	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Qy	448	AATTGCATGTATGGCGAAATCGTGGCCAACTTCGTCTCCAAAATTGCATGAAGG	507		
Db	1138	AATTGCATGTATGGCGAAATCGTGGCCAACTTCGTCTCCAAAATTGCATGAAGG	1197		
Oy	508	AAGGCTCTCCTAAGCAACATCGTTTCAAAGTTGGACAAGTCCTTGAACATAATTAATAT	567		

```
Db 1198 AAGCTCTCCTAAGCAAGCATCTTTCAAAAGTTGACAAAGCTCTTGAACTATTAATAT 1257
Oy 568 TCCAAATTCCTACTGAATAACGCGTAGCGGCAATTTCAAGAAATATGTGACGAGCAATGAT 627
Db 1258 TCCAAATTCCTACTGAATAACGCGTAGCGGCAATTTCAAGAAATATGTGACGAGCAATGAT 1317
Oy 628 GTATCTATCACAAGAAAGACTTTTCCGATGCTTCCAGATGCAAGATGACGACGACACA 687
Db 1318 GTATCTATCACAAGAAAGACTTTTCCGATGCTTCCAGATGCAAGATGACGACGACACA 1377
Oy 688 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAGGAGCTTCTTCATATTCTCTGT 747
Db 1378 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAGGAGCTTCTTCATATTCTCTGT 1437
Oy 748 GGATTTGACGAGCAATGATATCACTAAATGCGAAAGAAAGAAATATATTGACACACA 807
Db 1438 GGATTTGACGAGCAATGATATCACTAAATGCGAAAGAAAGAAATATATTGACACACA 1497
Oy 808 AATTAATATGCTCAAGCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 867
Db 1498 AATTAATATGCTCAAGCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 1557
Oy 868 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 927
Db 1558 ACATTTGATCAATTTAGCAAAAGATCCAAATTTGATCCCATGATTTGGAGAAAGTTAAGTT 1617
Oy 928 GGACAAAGTTTGAAGCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 987
Db 1618 GGACAAAGTTTGAAGCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 1677
Oy 988 TCGATTTCGAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1047
Db 1678 TCGATTTCGAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1737
Oy 1048 GCACCTTGAAGACAGTTTCCATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1107
Db 1738 GCACCTTGAAGACAGTTTCCATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1797
Oy 1108 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1857
Oy 1168 GATGAATACTTGGAGAAAGATCTGACAAACCTTACCGCTTGACTGTTTCAAGCAATG 1227
Db 1858 GATGAATACTTGGAGAAAGATCTGACAAACCTTACCGCTTGACTGTTTCAAGCAATG 1917
Oy 1228 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1280
Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1970

RESULT 7
US-09-872-523-78
; Sequence 78, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Eva M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 78
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-78
```

```
Query Match 55.6%; Score 833; DB 10; Length 2307;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 448 AATTCATGATTTGGGCAAAATGTCGGCCCAACTTCGGCTGCCAAATTCGATGAAGGG 507
Db 1138 AATTCATGATTTGGGCAAAATGTCGGCCCAACTTCGGCTGCCAAATTCGATGAAGGG 1197
Oy 508 AAGCTCTCTTAAGCAAGCATCTTTCAAGAGTTGGACAAAGCTTTGAACATTAATTAAT 567
Db 1198 AAGCTCTCTTAAGCAAGCATCTTTCAAGAGTTGGACAAAGCTTTGAACATTAATTAAT 1257
Oy 568 TCCAAATTCCTACTGAATAACGCGTAGCGGCAATTTCAAGAAATATGTGACGACGCAATGAT 627
Db 1258 TCCAAATTCCTACTGAATAACGCGTAGCGGCAATTTCAAGAAATATGTGACGACGCAATGAT 1317
Oy 628 GTATCTATCACAAGAAAGACTTTTCCGATGCTTCCAGATGCAAGATGACGACGACACA 687
Db 1318 GTATCTATCACAAGAAAGACTTTTCCGATGCTTCCAGATGCAAGATGACGACGACACA 1377
Oy 688 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAGGAGCTTCTTCATATTCTCTGT 747
Db 1378 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAGGAGCTTCTTCATATTCTCTGT 1437
Oy 748 GGATTTGACGAGCAATGATATCACTAAATGCGAAAGAAAGAAATATATTGACACACA 807
Db 1438 GGATTTGACGAGCAATGATATCACTAAATGCGAAAGAAAGAAATATATTGACACACA 1497
Oy 808 AATTAATATGCTCAAGCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 867
Db 1498 AATTAATATGCTCAAGCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 1557
Oy 868 ACATTTGATCAATTTAGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 927
Db 1558 ACATTTGATCAATTTAGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1617
Oy 928 GGACAAAGTTTGAAGCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 987
Db 1678 GGACAAAGTTTGAAGCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 1737
Oy 988 TCGATTTCGAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1047
Db 1678 TCGATTTCGAATTTTGGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1737
Oy 1048 GCACCTTGAAGACAGTTTCCATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1107
Db 1738 GCACCTTGAAGACAGTTTCCATTCATATCAATTAATACATTTATGTTCCCAAGTTGTTAT 1797
Oy 1108 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATATTTGGAACTTGTCCGCCAGATGAGTTCAAGAAACATTCAGATGG 1857
Oy 1168 GATGAATACTTGGAGAAAGATCTGACAAACCTTACCGCTTGACTGTTTCAAGCAATG 1227
Db 1858 GATGAATACTTGGAGAAAGATCTGACAAACCTTACCGCTTGACTGTTTCAAGCAATG 1917
Oy 1228 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1280
Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTTCTGATTTCCAAACGGGT 1970

RESULT 8
US-09-864-761-31987/C
; Sequence 31987, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aeomica-X-1
```

```

; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 31987
; LENGTH: 280
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009405.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EST HUMAN HIT: BE535426.1, EVALUE 2.00e-32
; OTHER INFORMATION: NT HIT: D84430.1, EVALUE 1.00e-32
; OTHER INFORMATION: SWISSPROT HIT: P15624, EVALUE 6.00e-08
US-09-864-761-31987

Query Match 1.3%; Score 20; DB 9; Length 280;
Best Local Similarity 100.0%; Pred. No. 8.3;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1387 GATGAGCAATTGATGAAC 1406
Db 156 GATGAGCAATTGATGAAC 137
```

```

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15473
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC009405.2
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
US-09-864-761-15473

Query Match 1.3%; Score 20; DB 9; Length 425;
Best Local Similarity 100.0%; Pred. No. 8.7;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1387 GATGAGCAATTGATGAAC 1406
Db 391 GATGAGCAATTGATGAAC 372
```

RESULT 10

```

US-10-136-224-2
; Sequence 2, Application US/10136224
; Publication No. US20020174448A1
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOCs)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
```

```

: TITLE OF INVENTION: IN HUMANS
: FILE REFERENCE: P07 42591
: CURRENT APPLICATION NUMBER: US/10/136, 224
: CURRENT FILING DATE: 2002-04-29
: PRIOR APPLICATION NUMBER: US/09/327, 138
: PRIOR FILING DATE: 1999-06-07
: NUMBER OF SEQ ID NOS: 30
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 2863
: TYPE: DNA
: ORGANISM: MURINE
: FEATURE:
: NAME/KEY: Promoter
: LOCATION: (-2759)...(104)
: US-10-136-224-2

```

Query Match	1.38;	Score 20;	DB 13;	Length 2863;
Best Local Similarity	100.0%;	Pred. No. 11;		
Matches 20;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	504	AGGSAAGGCTCTCCTAAGCA	523
Db	18	AGGSAAGGCTCTCCTAAGCA	37

```

RESULT 11
US-10-136-224-3
; Sequence 3, Application US/10136224
; Publication No. US2002017448A1
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC3)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591
; CURRENT APPLICATION NUMBER: US/10/136,224
; CURRENT FILING DATE: 2002-04-29
; PRIOR APPLICATION NUMBER: US/09/327,138
; PRIOR FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: MURINE
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (-2759)...(927)
; US-10-136-224-3

```

```
Query Match      1.3%; Score 20; DB 13; Length 3686;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0
```

Qy	504	AGGGAAGGCTCTCTAAGCA	523
Db	18	AGGGAAGGCTCTCTAAGCA	37

```

: RESULT 12
: US-10-136-224-1
: Sequence 1, Application US/10136224
: Publication No. US20020174448A1
: GENERAL INFORMATION:
: APPLICANT: AUERNHAMMER, CHRISTOPH J.
: APPLICANT: MELMED, SHLOMO
: TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
: TITLE OF INVENTION: (SOCs) -3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
: TITLE OF INVENTION: IN HUMANS
: FILE REFERENCE: P07 42591
: CURRENT APPLICATION NUMBER: US/10/136, 224

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

Query Match	1.3%;	Score 20;	DB 13;	Length 3940;
Best Local Similarity	100.0%;	Pred. No. 11;		
Matches	20;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY	504	AGGGAAGGCTCTCCTAAGCA	523
Db	166	AGGGAAGGCTCTCCTAAGCA	185

```

RESULT 13
US-10-240-485-46
Sequence 46, Application US/10240485
Publication No. US20030148327A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with
TITLE OF INVENTION: Metastasis
FILE REFERENCE: 5013.1007
CURRENT APPLICATION NUMBER: US/10/240,485
CURRENT FILING DATE: 2002-10-02
PRIOR APPLICATION NUMBER: PCT/EP01/03970
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 202
SEQ ID NO 46
LENGTH: 11996
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens
US-10-240-485-46

```

Query Match 1.3%; Score 20; DB 12; Length 11996;
Best Local Similarity 100.0%; Pred. No. 12;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 428 ATATGTAATTTATGTAAT 447
|||||
Db 10139 ATATGTAATTTATGTAAT 10158

RESULT 14
US-10-240-453-6/c
; Sequence 6, Application US/10240453
; Publication No. US20030148326A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPERBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
; TITLE OF INVENTION: Transcription
; TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
; FILE REFERENCE: 5013.1009
; CURRENT APPLICATION NUMBER: US/10/240,453
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03973
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 350
; SEQ ID NO 6
; LENGTH: 7450
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1729, 2292, 2294..2295, 2299, 2798)
US-10-240-453-6

Query Match 1.3%; Score 19; DB 12; Length 7450;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1073 ATATCAATATATACATTAT 1091
|||||
Db 6762 ATATCAATATATACATTAT 6744

RESULT 15
US-10-195-144-87
; Sequence 87, Application US/10195144
; Publication No. US20030126646A1
; GENERAL INFORMATION:
; APPLICANT: BROWN, GREGORY G.
; APPLICANT: FORMANOVA, NATASA
; APPLICANT: DENDY, CHARLES
; APPLICANT: LANDRY, BENJOIT S.
; APPLICANT: CHEUNG, WING
; APPLICANT: JIN, HUA
; TITLE OF INVENTION: NUCLEAR FERTILITY RESTORER GENES AND METHODS OF USE IN
; TITLE OF INVENTION: PLANTS
; FILE REFERENCE: 16313-0136
; CURRENT APPLICATION NUMBER: US/10/195,144
; CURRENT FILING DATE: 2002-10-01
; PRIOR APPLICATION NUMBER: 60/305,026
; PRIOR FILING DATE: 2001-07-12

; PRIOR APPLICATION NUMBER: 60/305,363
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/308,736
; PRIOR FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 87
; LENGTH: 271990
; TYPE: DNA
; ORGANISM: Raphanus sativum
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (144241)..(144300)
; OTHER INFORMATION: a, t, c, g, other or unknown
US-10-195-144-87

Query Match 1.3%; Score 19; DB 14; Length 271990;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1395 ATTTGATGACTGTATGAT 1413
|||||
Db 34401 ATTTGATGACTGTATGAT 34419

Search completed: September 2, 2003, 20:40:03
Job time : 409 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 15:09:29 ; Search time 454 Seconds
(without alignments)
8901.010 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atcgtcgaattcttgaat.....cgaaaagracactatga 1497

Scoring table: OLIGO NDC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_19jun03:*

1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1497	100.0	1497	24	ABL52497
2	1446	96.6	1497	24	ABL52515
3	1446	96.6	1497	24	ABL52516
4	1446	96.6	1497	24	ABL52517
5	833	55.6	2307	24	ABL52518
6	833	55.6	2307	24	ABL52519
7	833	55.6	2307	24	ABL52520
8	21	1.4	872	22	AAH04590

9	21	1.4	2202	22	AAH17401
10	20	1.3	280	22	ABA49041
11	20	1.3	280	22	AAI21886
12	20	1.3	280	22	AAI07575
13	20	1.3	280	22	ABSI15121
14	20	1.3	425	22	ABA43943
15	20	1.3	425	22	AAI12703
16	20	1.3	425	22	AAI02613
17	20	1.3	425	22	ABSO2626
18	20	1.3	263	22	MAC85357
19	20	1.3	3686	22	MAC83358
20	20	1.3	3940	22	MAC85356
21	20	1.3	11996	24	ABL34493
22	22	1.3	2127	25	ABD51265
23	19	1.3	2178	25	ABX06430
24	19	1.3	2181	25	ABE42211
25	19	1.3	2883	21	AC42886
26	19	1.3	3262	23	ABL17787
27	19	1.3	6870	23	ABL17786
28	19	1.3	7450	24	ABK28132
29	19	1.3	8494	19	AAV52296
30	19	1.3	8753	24	ABL92270
31	19	1.3	35048	23	ABU06102
32	19	1.3	2162598	25	ABS56454
33	18	1.2	108	17	AAI12617
34	18	1.2	286	24	ABX92085
35	18	1.2	302	24	AD37370
36	18	1.2	360	24	AB332053
37	18	1.2	367	22	AK76409
38	18	1.2	409	22	ABA46491
39	18	1.2	512	24	ABQ57224
40	18	1.2	527	24	ABN76926
41	18	1.2	539	23	ABV54427
42	18	1.2	742	21	AAZ98082
43	18	1.2	742	22	ABD11695
44	18	1.2	742	24	ABK69791
45	18	1.2	781	21	AA297460

ALIGNMENTS

RESULT 1
ABL52497
ABL52497 standard; DNA: 1497 BP.

XX AC ABL52497;
XX DT 17-JUL-2002 (first entry)
XX DE Caenorhabditis elegans lin-61 nucleotide sequence SEQ ID NO:6.
XX DE
XX XX Caenorhabditis elegans; lin-8; lin-56; lin-61: tumour suppressor;
XX KW cell proliferation; nematode; cancer; gene; ds.
XX OS Caenorhabditis elegans.
XX FH
XX FT 1.1497 Location/Qualifiers
XX FT CDS /tag= a
XX FT /product= "LIN-61 protein"
XX PD
XX PN W0200194545-A2.
XX PD 13-DEC-2001.
XX PF 01-JUN-2001: 2001WO-US17909.
XX PR 02-JUN-2000: 2000US-208802P.
XX PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX PI Horvitz HR, Davison EM, Lu X;

XX WPI, 2002-401590/43.
DR P-PSDB: ABB78650.
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Claim 24; Page 74; 116pp; English.
XX
XX The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring LIN-8,
CC LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
CC using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes C. elegans
CC LIN-61 from the present invention.
XX
XX Sequence 1497 BP; 497 A; 272 C; 312 G; 416 T; 0 other:
SQ
Query Match 100.0%; Score 1497; DB 24; Length 1497;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 541 GGCACAGCTCTTGACATTTAATTTCCATTTCTAGTAATACGCGTACGGCGAATT 600
DB 541 GGACACAGCTCTTGACATTTAATTTCCATTTCTAGTAATACGCGCGAATT 600
QY 601 CAAAGAAATATGTGACGACGAAATGATATCTATCAAGAAAGAAAGACTTCCGATGCG 660
DB 601 CAAAGAAATATGTGACGACGAAATGATATCTATCAAGAAAGAAAGACTTCCGATGCG 660
QY 661 CTTCAGATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
DB 661 CTTCAGATGACATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
QY 721 GAGGAGAGCTTCTCATATTTTCCCTGTTGATTTGACAGATGATGATGATGATGATGAT 780
DB 721 GAGGAGAGCTTCTCATATTTTCCCTGTTGATTTGACAGATGATGATGATGATGATGAT 780
QY 781 GCGAAAAAGAAATATTTGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
DB 781 GCGAAAAAGAAATATTTGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
QY 841 AATCCAGATATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
DB 841 AATCCAGATATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
QY 901 CCCATGATTTGAGAGAAAGTTAAGTTGGAGAAAGTTGAGCTCATGACCCCTTGAGCT 960
DB 901 CCCATGATTTGAGAGAAAGTTAAGTTGGAGAAAGTTGAGCTCATGACCCCTTGAGCT 960
QY 961 CAGCAATTCATTAACCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 CAGCAATTCATTAACCTCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 CTATATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
DB 1021 CTATATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
QY 1081 AATCATTTATGTTCCGAGTGGTATGCGGAAAGTATATTTGGAACCTGTTCCGCCCA 1140
DB 1081 AATCATTTATGTTCCGAGTGGTATGCGGAAAGTATATTTGGAACCTGTTCCGCCCA 1140
QY 1141 GATGAGTTCAAGAAAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
DB 1141 GATGAGTTCAAGAAAGAAATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1200
QY 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGAGATTTAAGGTA 1260
DB 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGAGATTTAAGGTA 1260
QY 1261 AATCGATTTCCAAAGCGGTAGGACTAGCGCTGAGCTGCTGACATGTTGAAATATCAG 1320
DB 1261 AATCGATTTCCAAAGCGGTAGGACTAGCGCTGAGCTGCTGACATGTTGAAATATCAG 1320
QY 1321 TTTATTTGTCAGCTACAGTAAGTAATCACTTCATGGAAGATGATGATGATGATGATGATGAT 1380
DB 1321 TTTATTTGTCAGCTACAGTAAGTAATCACTTCATGGAAGATGATGATGATGATGATGATGAT 1380
QY 1381 GCGTGGATGAGAAATTTGATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
DB 1381 GCGTGGATGAGAAATTTGATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1440
QY 1441 GGATGCTGTGAAGCGCAAGTATGTTCTACAACCTCCGAAAGTCAACTATGTA 1497
DB 1441 GGATGCTGTGAAGCGCAAGTATGTTCTACAACCTCCGAAAGTCAACTATGTA 1497

RESULT 2
ABL52515
ID ABL52515 standard; DNA: 1497 BP.
XX ABL52515;
AC
XX
XX 17-JUL-2002 (first entry)

XX DE C. elegans mutant lin-61 DNA sequence lin-61(n3446) SEQ ID NO:73.
 XX XX
 XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
 KW cell proliferation; nematode; cancer; mutant; gene; ds.
 XX
 OS Caenorhabditis elegans.
 XX
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1497
 FT /tag- a
 FT /product= "LIN-61(n3446) protein"
 XX
 PN WO200194545-A2.
 XX
 PD 13-DEC-2001.
 XX
 XX 01-JUN-2001; 2001WO-US17909.
 XX
 XX 02-JUN-2000; 2000US-208802P.
 XX
 XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
 XX
 XX Horvitz HR, Davidson EM, Lu X;
 PI WPI; 2002-401590/43.
 XX
 DR P-PSDB; ABB78697.
 XX
 XX
 PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
 PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
 PT diagnosing and treating cell proliferative diseases such as cancer
 XX
 PS Claim 27: Page 112-113; 116pp; English.
 XX
 CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
 CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
 CC cytostatic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that
 CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
 CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumor suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence encodes a
 CC C. elegans mutant LIN-61 protein from the present invention.
 XX
 XX Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;

Query Match 96.6%; Score 1446; DB 24; Length 1497;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGTGGAATTCGAAAATGTGAGCTACAAAATGCGACAGAAACTCGATAG 60
 DB 1 ATGTGGAATTCGAAAATGTGAGCTACAAAATGCGACAGAAACTCGATAG 60
 QY 61 ACCTACTGTGGGAATCCATTTACATCAGTTCGAGAAAGAAACTCTTCAATTC 120
 DB 61 ACCTACTGTGGGAATCCATTTACATCAGTTCGAGAAAGAAACTCTTCAATTC 120
 QY 121 GTTGAAGCATTCATCGTACGTTACGTTAAATTTAAAGAAATGCGTGAAGAGAGTT 180
 DB 121 GTTGAAGCATTCATCGTACGTTACGTTAAATTTAAAGAAATGCGTGAAGAGAGTT 180

QY 181 ATCTTGAAGACAGTGTCCATGATTATGACAAAGACTGGATTCGATTCAGATGAG 240
 DB 181 ATCTTGAAGACAGTGTCCATGATTATGACAAAGACTGGATTCGATTCAGATGAG 240
 QY 241 TTTCGACGAAATGAAAAGTTTGGCGATACAGAGTTTGGCTCAGTTTATGGAGCTGAC 300
 DB 241 TTTCGACGAAATGAAAAGTTTGGCGATACAGAGTTTGGCTCAGTTTATGGAGCTGAC 300
 QY 301 ACGAAATTTGGCTCAATATTTTATGCGACGATATGTTGGTTGGCAAGCCGCAATG 360
 DB 301 ACGAAATTTGGCTCAATATTTTATGCGACGATATGTTGGTTGGCAAGCCGCAATG 360
 QY 361 AGTATCCCAATATGATGATTAATGATGCTCCGCGCTTCGCAATCAGCAAGATATC 420
 DB 361 AGTATCCCAATATGATGATTAATGATGCTCCGCGCTTCGCAATCAGCAAGATATC 420
 QY 421 CAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 DB 421 CAAATGATATGTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 480
 QY 481 TGGCTGTCTCCAAATTTGATGAAGGAGGCTCTCTTAAGCAAGCATCTTCAAGTT 540
 DB 481 TGGCTGTCTCCAAATTTGATGAAGGAGGCTCTCTTAAGCAAGCATCTTCAAGTT 540
 QY 541 GGACAACGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 DB 541 GGACAACGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 600
 QY 601 CAGAAATATGTCGACGACGAAATGATATATACAAAGAAAGCTTCCCAATCG 660
 DB 601 CAGAAATATGTCGACGACGAAATGATATATACAAAGAAAGCTTCCCAATCG 660
 QY 661 CTTCCAGATGCAATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
 DB 661 CTTCCAGATGCAATGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
 QY 721 GAGGGAAGCTTCTTCAATATTTCTGTTGGATTGGACAGTCAATGATATCACTAAT 780
 DB 721 GAGGGAAGCTTCTTCAATATTTCTGTTGGATTGGACAGTCAATGATATCACTAAT 780
 QY 781 GCGAAAAGCAATATATGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
 DB 781 GCGAAAAGCAATATATGACGACGACGACGACGACGACGACGACGACGACGACGAC 840
 QY 841 AATCCAAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
 DB 841 AATCCAAAGTATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
 QY 901 CCCATGATTTGGAGAAAGTTAAGTTGGACAAAGTTGAGCTCATGACCCCTTGCT 960
 DB 901 CCCATGATTTGGAGAAAGTTAAGTTGGACAAAGTTGAGCTCATGACCCCTTGCT 960
 QY 961 CAGCAATTCATTAACCTCAGCTGCTGATTCGCAATTTTGCAGAACTAAGAGAT 1020
 DB 961 CAGCAATTCATTAACCTCAGCTGCTGATTCGCAATTTTGCAGAACTAAGAGAT 1020
 QY 1021 CTTATTTGGGAATGAGTGTGCTCAGATGCTGAGAGAGCTTCTATTTATATCAAT 1080
 DB 1021 CTTATTTGGGAATGAGTGTGCTCAGATGCTGAGAGAGCTTCTATTTATATCAAT 1080
 QY 1081 AATACATTTATGTTCCAGATGCTGATGCGGAAAGTAAATTTGCAACTTTCCGCA 1140
 DB 1081 AATACATTTATGTTCCAGATGCTGATGCGGAAAGTAAATTTGCAACTTTCCGCA 1140
 QY 1141 GATGAGTTCAAGGAACATTCAGATGCTGATGCGGAAAGTAAATTTGCAACTTTCCGCA 1200
 DB 1141 GATGAGTTCAAGGAACATTCAGATGCTGATGCGGAAAGTAAATTTGCAACTTTCCGCA 1200
 QY 1201 CTACCGCTTGAATTTGCAAGCAATGCTTCCCAAGAGATTAGCAAAATTTAAGTA 1260
 DB 1201 CTACCGCTTGAATTTGCAAGCAATGCTTCCCAAGAGATTAGCAAAATTTAAGTA 1260

Oy		1261	ATTTGTGATTTCCTCAACGGGTAGGACTACAGTCCCTTGAAAGCGTGTGCATCTGTGAATAATCAG	1320
Db		1261	TTTTTTTTCATTTCCAAACGGGTAGGACTACGCCCTTGAAAGCGTGTGCATCTGTGAATAATCAG	1320
Oy		1321	TTTTATTGCTCAGCTACAGTAAGTAATTCAGTTCATGGAAGACATGATMAATGTCAAATTTCCGAC	1380
Db		1321	TTTTTTTGTTCGACGCTACAGTAAGTAATTCAGTTCATGGAAGACATGATMAATGTCAAATTTCCGAC	1380
Oy		1381	GGCTGGGATGAGAAGATTGTGATGAACCTGTATGATGTGGACTCCCATGATATTTCCACCATA	1440
Db		1381	GGCTGGGATGAGAAGATTGTGATGAACCTGTATGATGTGGACTCCCATGATATTTCTACCCATA	1440
Oy		1441	GGATGGTGTGAAGGCGCACGTTATGTTCTACAACTCCGAAAAGATPACAACTATTGA	1497
Db		1441	GGATGGTGTGAAGGCGCACGTTATGTTCTACAACTCCGAAAAGATPACAACTATTGA	1497
RESULT 3				
ABL52516		ID	ABL52516 standard; DNA; 1497 bp.	
XX		AC	ABL52516;	
XX		DT	17-JUN-2002 (first entry)	
XX		DE	C. elegans mutant lin-61 DNA sequence lin-61(n3447) SEQ ID NO:74.	
XX		KW	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;	
XX		KM	cell proliferation; nematode; cancer; mutant; gene; ds.	
XX		OS	Caenorhabditis elegans.	
XX		FH	Key Location/Qualifiers	
FT		CDS	1..1497	
FT			/tag= a	
FT			/product= "LIN-61(n3447) protein"	
PX			WO200194545-A2.	
PN			13-DEC-2001.	
PD			01-JUN-2001; 2001WO-US17909.	
PE			02-JUN-2000; 2000US-208802P.	
PR			(MASI) MASSACHUSETTS INST TECHNOLOGY.	
PA			HORVITZ HR, DAVISON EM, LU X;	
PI			WPt: 2002-401590/43.	
DR			P-PStDb; ABB78698.	
XX			New Caenorhabditis elegans nucleic acid involved in tumor suppressor	
PT			pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for	
PP			diagnosing and treating cell proliferative diseases such as cancer -	
PS			Claim 27; Page 113; 116pp; English.	
XX			The present invention describes a substantially pure nematode C. elegans	
CC			(Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide	
CC			(see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide	
CC			(see ABB78650), where the polypeptides comprise at least 130, 110, 130	
CC			contiguous amino acids of the 386, 322 and 498 amino acid sequences given	
CC			in ABB78648 to ABB78650, and modulate cell proliferation. (I) has	
CC			cytostatic activity, and can be used in gene therapy. (I) is useful for	
CC			modulating proliferation of a cell, and for identifying the compound that	
CC			modulates cell proliferation. (I) can be used for diagnosing an animal	
CC			(preferably, human) for the presence of the cell proliferation disease,	
CC			or an increased chance of developing the disease, by measuring lin-8,	
CC			lin-56 or lin-61 nucleic acid expression in a sample obtained from the	
CC			animal, by measuring amount of lin-8, lin-56 or lin-61 polypeptide in the sample	
CC			(I) is also useful for identifying the compound that modulates cell	
CC			proliferation.	

CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC *C. elegans* mutant LTN-61 protein from the present invention.
XX
SQ Sequence 1497 BP; 498 A; 272 C; 311 G; 416 T; 0 other;

Query Match	96.68;	Score 1446;	DB 24;	Length 1497;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 1496; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

QY	1	TTGCTCGAAATTCGAAAATGTGAGAGCTAACAAAAATCGCAGAAAACTGTGATAG	60
Db	1	ATGCTCGAATTTCTGAAAATGTGTGAGAGCTAACAAAAATCGCAGAAAACTGTGATAG	60
QY	61	ACCTACTGTGGGAAATCCATTTATCATCAGTTGCGAGAAAAGAAAACCTCTTTCATTTCA	120
Db	61	ACCTACTGTGGGAAATCCATTTATCATCAGTTGCGAGAAAAGAAAACCTCTTTCATTTCA	120
QY	121	GTTGAGATTCATTCGTAACCTTACGTTAAATTTTAAAGAAATCGTGAAGGAGGTT	180
Db	121	GTTGAGATTCATTCGTAACCTTACGTTAAATTTTAAAGAAATCGTGAAGGAGGTT	180
QY	181	ATCTTCGAAACAGTGTGTCATGATTATGACAAAGACGCAATGATTCAGATCAGATGG	240
Db	181	ATCTTCGAAACAGTGTGTCATGATTATGACAAAGACGCAATGATTCAGATCAGATGG	240
QY	241	TTTGACAGAAATGAAAAAGTTTGGCGGATACAGAGCTCTG6CTCAGTTTATCGAGCTGAC	300
Db	241	TTTGACAGAAATGAAAAAGTTTGGCGGATACAGAGCTCTG6CTCAGTTTATCGAGCTGAC	300
QY	301	ACGAAATTTGGCTCAATATTTATTCGAGAGATATGTTGGTTGGCAACGCCGCAATG	360
Db	301	ACGAAATTTGGCTCAATATTTATTCGAGAGATATGTTGGTTGGCAACGCCGCAATG	360
QY	361	AGTGATCCCAATATGATGATAAATTTGTATATCTCCGCGCTTGCAATCAACGAAATAC	420
Db	361	AGTGATCCCAATATGATGATAAATTTGTATATCTCCGCGCTTGCAATCAACGAAATAC	420
QY	421	CAAAATGATATGTAATATATGTAATATATTCATTCATTCGCGCAAAATCGTGGCGAACT	480
Db	421	CAAAATGATATGTAATATATGTAATATATTCATTCATTCGCGCAAAATCGTGGCGAACT	480
QY	481	TGCGGTGTCCAAAATTTGATGAAGGGAGGCTCCCTAACCAAGCATCGTTCCAAAGTT	540
Db	481	TGCGGTGTCCAAAATTTGATGAAGGGAGGCTCCCTAACCAAGCATCGTTCCAAAGTT	540
QY	541	GGACAAGCTCTGGAATTTTAAATTTATTCCAATTCCTACTGAAATACGGGTAGCGCAAT	600
Db	541	GGACAAGCTCTGGAATTTTAAATTTATTCCAATTCCTACTGAAATACGGGTAGCGCAAT	600
QY	601	CAGCAATATGTGGACGACGCAATGATATCTATCACAAAGAAAAGCTTCCCGAATCG	660
Db	601	CAGCAATATGTGTGGACGACGCAATGATATCTATCACAAAGAAAAGCTTCCCGAATCG	660
QY	661	CTTCCAGATGAGATGACGACGACAAGTCTTATGCTGTGATCTCAATATTTGGATAGAC	720
Db	661	CTTCCAGATGAGATGACGACGACAAGTCTTATGCTGTGATCTCAATATTTGGATAGAC	720
QY	721	GAGGAAACCTCTTCATATTTCTGTTGGATTTTGACGACGTCATGATATCACTAAAT	780
Db	721	GAGGAAACCTCTTCATATTTCTGTTGGATTTTGACGACGTCATGATATCACTAAAT	780
QY	781	GCGAAAAAGAAATATATTGACGACACCAATATTAATGCTCAAGCAATATAAAAATGGAGAA	840
Db	781	GCGAAAAAGAAATATATTGACGACACCAATATTAATGCTCAAGCAATATAAAAATGGAGAA	840
QY	841	AATCCAAATATGACTGACGACGACCTCATTGTGATCAATTAAGCAAAAGATCCAAATGAT	900
Db	841	AATCCAAATATGACTGACGACGACCTCATTGTGATCAATTAAGCAAAAGATCCAAATGAT	900

QY	901	CCAGTGAATTTGGAAAAAGTTAAAGTTGGAGCAAAAAGTTTGGCTCATCGACCCCTGGCT	960
Db	901	CCCATGATTTGGAAAAAGTTAAAGTTGGAGCAAAAAGTTTGGCTCATCGACCCCTGGCT	960
QY	961	CAGCAATTCATATAACCTCCACGTCGCTGCATTCCTCAAAATTTTGCAAAACGGAAGATAT	1020
Db	961	CAGCAATTCATATAACCTCCACGTCGCTGCATTCCTCAAAATTTTGCAAAACGGAAGATAT	1020
QY	1021	CTTATTTGGGAATGATGGTCCAGATGCACTTGAAGACAGTTTCTATTCATATCAAT	1080
Db	1021	CTTATTTGGGAATGATGGTCCAGATGCACTTGAAGACAGTTTCTATTCATATCAAT	1080
QY	1081	AATACATTTTATGTCCAGTTGGTTATGGGAAAAAGTATATTTGSAACCTGTCCGGCA	1140
Db	1081	AATACATTTTATGTCCAGTTGGTTATGGGAAAAAGTATATTTGSAACCTGTCCGGCA	1140
QY	1141	GATGAGTTCAAAGAGACATTCAGATGGGATGAATACTTGGAGAAAGAAATCGCAGAAAC	1200
Db	1141	GATGAGTTCAAAGAGACATTCAGATGGGATGAATACTTGGAGAAAGAAATCGCAGAAAC	1200
QY	1201	CTACCGCTTGACTTGTTCACAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGSTA	1260
Db	1201	CTACCGCTTGACTTGTTCACAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGSTA	1260
QY	1261	ATTCTGATTTCCAAACGGGTAGAGACTACGCCCTTGAAGTGGTGACATGTGGAAAATCAG	1320
Db	1261	ATTCTGATTTCCAAACGGGTAGAGACTACGCCCTTGAAGTGGTGACATGTGGAAAATCAG	1320
QY	1321	TTTATTTGTGCAGCTACAGTCAATTCATGCAAGACTGATTAATATGTTCAATTTGCAC	1380
Db	1321	TTTATTTGTGCAGCTACAGTCAATTCATGCAAGACTGATTAATATGTTCAATTTGCAC	1380
QY	1381	GGCTGGAGTGAAGAATTTGATGAACCTGTATGATGTGGACCTCCATGATATTTCTACCGATA	1440
Db	1381	GGCTGGAGTGAAGAATTTGATGAACCTGTATGATGTGGACCTCCATGATATTTCTACCGATA	1440
QY	1441	GGATGGGTGTAAGCGCACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATATGA	1497
Db	1441	GGATGGGTGTAAGCGCACAGTTATGTTCTACAACTCCGAAAAAGTACAACTATATGA	1497

RESULT 4	ABL52517	ABL52517 standard; DNA; 1497 BP.
ID	ABL52517	
XX	ABL52517;	
AC		
XX		
DT	17-JUL-2002 (first entry)	
XX		
XX		
DE	C. elegans mutant lin-61 DNA sequence lin-61(n3624) SEQ ID NO:75.	
XX		
KW	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;	
XX	cell proliferation; nematode; cancer; mutant; gene; ds.	
OS	Caenorhabditis elegans.	
XX		
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1497
FT		/*tag= a
FT		/product= "LIN-61(n3624) protein"
XX		
XX	WO200194545-A2.	
XX		
PD	13-DEC-2001.	
XX		
XX		
PF	01-JUN-2001; 2001WO-US17909.	
XX		
XX		
PR	02-JUN-2000; 2000US-208802P.	
XX		
PA	(MAST) MASSACHUSETTS INST TECHNOLOGY.	
XX		
PI	Horvitz HR, Davison EM, Lu X;	
XX		

DR WPI: 2002-401590/43.
DR P-PSDB: ABB78699.
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PI pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PI diagnosing and treating cell proliferative diseases such as cancer -
XX
PS Claim 27, page 113-114; 116pp; English.

C. elegans present invention describes a substantially pure nematode C. elegans (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-61 polypeptide (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide (see ABB78650), where the polypeptides comprise at least 130, 110, 130 contiguous amino acids of the 386, 322 and 498 amino acid sequences given in ABB78648 to ABB78650, and modulate cell proliferation. (I) has cytosolic activity, and can be used in gene therapy. (I) is useful for modulating proliferation of a cell, and for identifying the compound that modulates cell proliferation. (I) can be used for diagnosing an animal (preferably, human) for the presence of the cell proliferation disease, or an increased chance of developing the disease, by measuring lin-8, lin-56 or lin-61 nucleic acid expression in a sample obtained from the animal, by measuring lin-8, lin-56 or lin-61 mRNA in the sample using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample. (I) is also useful for identifying the compound that modulates cell proliferation. (I) and its mutants are useful for understanding cell proliferative diseases including cancer, as well as in diagnosing and treating cell proliferative diseases. (I) is also useful for identifying tumour suppressors in other species such as mammals and may be used to identify therapeutic compounds. The present sequence encodes a C. elegans mutant LIN-61 protein from the present invention.

Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;

Query Match	96.68;	Score 1446;	DB 24;	Length 1497;
Best Local Similarity	99.98;	Pred. No. 0;		
Matches 1496; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0;

OY	1	ATGTCGAAATTTTCGAAAAATGTCAGAGCTAAACAAAAATGCGACAGAAAACTCGATTAG	60
Db	1	ATGTCGAAATTTTCGAAAAATGTCAGAGCTAAACAAAAATGCGACAGAAAACTCGATTAG	60
OY	61	ACCTACTGTGGGAATCCTATTACACAGTTCGAGAAAGGAAAACTCTTTATTCCA	120
Db	61	ACCTACTGTGGGAATCCTATTACACAGTTCGAGAAAGGAAAACTCTTTATTCCA	120
OY	121	GTTGAAGCATTCAATCGTAACTTACAGTTAATTTTAAACGAATGCGTGAAGAGGATT	180
Db	121	GTTGAAGCATTCAATCGTAACTTACAGTTAATTTTAAACGAATGCGTGAAGAGGATT	180
OY	181	ATCTCGAAACAGTGTGCTCATGATTATGACAAAGAACTGCAATTCGATTCAAGTCAGATGG	240
Db	181	ATCTCGAAACAGTGTGCTCATGATTATGACAAAGAACTGCAATTCGATTCAAGTCAGATGG	240
OY	241	TTTGACACGAATTAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGAC	300
Db	241	TTTGACACGAATTAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGAC	300
OY	301	ACGAAATTTTGGCTCAATTTTATTCGGACGATATGTTGGTTGGTGGCAAGCGCGCAATG	360
Db	301	ACGAAATTTTGGCTCAATTTTATTCGGACGATATGTTGGTTGGTGGCAAGCGCGCAATG	360
OY	361	AGTCATCCCAATATGATATAAATTTGTATATGCTCCGCGCTTGCAATACAGGAAGATTAC	420
Db	361	AGTCATCCCAATATGATATAAATTTGTATATGCTCCGCGCTTGCAATACAGGAAGATTAC	420
OY	421	CAAAATGATATGCTAAATTAATGTAATAATTGCATTTGATGGCGAAATGTCGGCCAAACT	480
Db	421	CAAAATGATATGCTAAATTAATGTAATAATTGCATTTGATGGCGAAATGTCGGCCAAACT	480
OY	481	TGCGCTGCTCCAAAATTCGATGAAGGAAGGCTCTCCATACGACGATCGTTTCAAAGTT	540
Db	481	TGCGCTGCTCCAAAATTCGATGAAGGAAGGCTCTCCATACGACGATCGTTTCAAAGTT	540

```

OY 541 GGACAAAGCTGTGAACTATTAAATTAATTCCTCAATTCCTGTAATACCGCGGAATT 600
    |||
    |||
    |||
Db 541 GGACAAAGCTGTGAACTATTAAATTAATTCCTCAATTCCTGTAATACCGCGGAATT 600
OY 601 CAAGAAATATGTGAGACGACGAATGAATGATCTATCAACAAAGAAAGACTTCCGGAATG 660
    |||
    |||
    |||
Db 601 CAAGAAATATGTGAGACGACGAATGAATGATCTATCAACAAAGAAAGACTTCCGGAATG 660
OY 661 CTTCAGATGACAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
    |||
    |||
    |||
Db 661 CTTCAGATGACAGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
OY 721 GAGGGAAGCTTCTTCATATTTCTCTGTTGATTTCCACAGTCAATGCAATCAATCAAT 780
    |||
    |||
    |||
Db 721 GAGGGAAGCTTCTTCATATTTCTCTGTTGATTTCCACAGTCAATGCAATCAATCAAT 780
OY 781 GCGAAAAAGGAATATATTTGAGACACACAAATTAATTTGCTCAAGCAATTAATTAATGAGAA 840
    |||
    |||
    |||
Db 781 GCGAAAAAGGAATATATTTGAGACACACAAATTAATTTGCTCAAGCAATTAATTAATGAGAA 840
OY 841 AATCAAGATATGACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
    |||
    |||
    |||
Db 841 AATCAAGATATGACTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 900
OY 901 CCCATGATTTGAGAAAAAGTTAAGTTGAGCAAAAGTTGAGCTCATCGACCCCTGGCT 960
    |||
    |||
    |||
Db 901 CCCATGATTTGAGAAAAAGTTAAGTTGAGCAAAAGTTGAGCTCATCGACCCCTGGCT 960
OY 961 CAGCAATTCATATACCTCCACAGTGCCTGCATTCCTCAATTTTCCAAACCTGAGAGAT 1020
    |||
    |||
    |||
Db 961 CAGCAATTCATATACCTCCACAGTGCCTGCATTCCTCAATTTTCCAAACCTGAGAGAT 1020
OY 1021 CTATATGTTGGAATGATGTTGCTCAGATGACCTTGAAGACGTTTCTATTCATATCAAT 1080
    |||
    |||
    |||
Db 1021 CTATATGTTGGAATGATGTTGCTCAGATGACCTTGAAGACGTTTCTATTCATATCAAT 1080
OY 1081 AATCATTTATGTTCCCAAGTGTGTTAGCGGAAAAAGTTAATTTGGAACCTTGTCCGCA 1140
    |||
    |||
    |||
Db 1081 AATCATTTATGTTCCCAAGTGTGTTAGCGGAAAAAGTTAATTTGGAACCTTGTCCGCA 1140
OY 1141 GATGAGTTCAAGAAACATTCAGATGGGATGAATTAATTCAGAAAGAAATTCGACGAAACC 1200
    |||
    |||
    |||
Db 1141 GATGAGTTCAAGAAACATTCAGATGGGATGAATTAATTCAGAAAGAAATTCGACGAAACC 1200
OY 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGAGATTAAGCAAAATTTAAAGTA 1260
    |||
    |||
    |||
Db 1201 CTACCGCTTGACTGTTCAACGCAATGCTTCCCAAGAGATTAAGCAAAATTTAAAGTA 1260
OY 1261 AATCGATTTCCAAAGCGGTAGGACTAGCCTTGAGCTGCTGACATGTGGAATATGAG 1320
    |||
    |||
    |||
Db 1261 AATCGATTTCCAAAGCGGTAGGACTAGCCTTGAGCTGCTGACATGTGGAATATGAG 1320
OY 1321 TTATATTTGTCAGATGACGTAATTCAGTTTCATGGAAGACGATTAATTCATTTTCGAC 1380
    |||
    |||
    |||
Db 1321 TTATATTTGTCAGATGACGTAATTCAGTTTCATGGAAGACGATTAATTCATTTTCGAC 1380
OY 1381 GCGTGGGATGAAGAAATTTGATGAACCTGATGATGGAATGGAATTTTACCGATA 1440
    |||
    |||
    |||
Db 1381 GCGTGGGATGAAGAAATTTGATGAACCTGATGATGGAATGGAATTTTACCGATA 1440
OY 1441 GGATGGTGTGAAGGCGACATTAATTTTACAACTCCGAAAAAGTAAACATATTA 1497
    |||
    |||
    |||
Db 1441 GGATGGTGTGAAGGCGACATTAATTTTACAACTCCGAAAAAGTAAACATATTA 1497

```

```

DE C. elegans lin-61 DNA sequence SEQ ID NO:76.
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX
OS Caenorhabditis elegans.
XX
PN W0200194545-A2.
XX
PD 13-DEC-2001.
XX
PE 01-JUN-2001; 2001WO-US1909.
XX
PR 02-JUN-2000; 2000US-208802P.
XX
PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Horvitz HR, Davison EM, Lu X;
XX
DR WPI: 2002-401590/43.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Disclosure; Page 114-115; 116pp; English.
XX
XX
XX The present invention describes a substantially pure nematode C. elegans
XX (Caenorhabditis elegans) nucleic acid (1) encoding a LIN-8 polypeptide
XX (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
XX (see ABB78650), where the polypeptides comprise at least 130, 110, 130
XX contiguous amino acids of the 386, 322 and 498 amino acid sequences given
XX in ABB78648 to ABB78650, and modulate cell proliferation. (1) has
XX cytosolic activity, and can be used in gene therapy. (1) is useful for
XX modulating proliferation of a cell, and for identifying the compound that
XX modulates cell proliferation. (1) can be used for diagnosing an animal
XX (preferably, human) for the presence of the cell proliferation disease,
XX or an increased chance of developing the disease, by measuring lin-8,
XX lin-56 or lin-61 nucleic acid expression in a sample obtained from the
XX animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
XX using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
XX (1) is also useful for identifying the compound that modulates cell
XX proliferation. (1) and its mutants are useful for understanding cell
XX proliferative diseases including cancer, as well as in diagnosing and
XX treating cell proliferative diseases. (1) is also useful for identifying
XX tumour suppressors in other species such as mammals and may be used to
XX identify therapeutic compounds. The present sequence represents a
XX C. elegans lin-61 DNA sequence from the present invention.
XX
XX Sequence 2307 BP; 740 A; 393 C; 466 G; 708 T; 0 other;
XX
XX
XX Query Match 55.6%; Score 833; DB 24; Length 2307;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 448 AATTCATGATGATGGGAAATGCTGGCCAACTGCTGCTCCAAATTCGATGAAGG 507
    |||
    |||
    |||
Db 1138 AATTCATGATGATGGGAAATGCTGGCCAACTGCTGCTCCAAATTCGATGAAGG 1197
OY 508 AAGGCTCTCTAGCAAGCATCGTTCAAGTTGGACAACGCTTGAACATAATTAAT 567
    |||
    |||
    |||
Db 1198 AAGGCTCTCTAGCAAGCATCGTTCAAGTTGGACAACGCTTGAACATAATTAAT 1257
OY 568 TCCAATTTCTAGTAATACGGGTAGCGGCAATTCAGAAATATGTGACGACGATGAT 627
    |||
    |||
    |||
Db 1258 TCCAATTTCTAGTAATACGGGTAGCGGCAATTCAGAAATATGTGACGACGATGAT 1317
OY 628 GTATCTATCACAAGAAAGACCTCCGCAATCGCTCCAGATGACGATGACGACGACAA 687
    |||
    |||
    |||
Db 1318 GTATCTATCACAAGAAAGACCTCCGCAATCGCTCCAGATGACGATGACGACGACAA 1377
OY 688 GTCCTTACCTGATGTCATATTTGATTAACGAGGAGGACCTTCTCATATTTCTGTT 747
    |||
    |||
    |||

```

RESULT 5
ID ABL52518 standard; DNA; 2307 BP.

ABL52518;

XX 17-JUL-2002 (first entry)

Db 1378 GTCCTTAGCTCTGATCTCATATATGATAGACGAGGAACTTCTTCAATATTTCCGTGTT 1437
 QY 748 GGATTTGACAGCATGATGATATCACTAAATGCGAAAAAGATATTTATGACACACA 807
 Db 1438 GGATTTGACAGCATGATGATATCACTAAATGCGAAAAAGATATTTATGACACACA 1497
 QY 808 AATTAATTTGCTCAAGCAATTAATAATGAGAAAAATCCAGATATGACTCAGACAGCTC 867
 Db 1498 AATTAATTTGCTCAAGCAATTAATAATGAGAAAAATCCAGATATGACTCAGACAGCTC 1557
 QY 868 ACATTTATCATATTAAGCAAAAAGATCAATTCATCCATGATTTGGAAAAAGTTAAGCTT 927
 Db 1558 ACATTTATCATATTAAGCAAAAAGATCAATTCATCCATGATTTGGAAAAAGTTAAGCTT 1617
 QY 928 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 987
 Db 1618 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 1677
 QY 988 TCGATTTCTCAAAATTTTGCAAAACCTGAAGAGATATCTTATTTGGGAATGATGTCAGAT 1047
 Db 1678 TCGATTTCTCAAAATTTTGCAAAACCTGAAGAGATATCTTATTTGGGAATGATGTCAGAT 1737
 QY 1048 GCATTTGAAGCAATTTTCTCATATCATATCAATTAATACATTTATGTTCCAGTTGTTAT 1107
 Db 1738 GCATTTGAAGCAATTTTCTCATATCATATCAATTAATACATTTATGTTCCAGTTGTTAT 1797
 QY 1108 GCGGAAAAGTATATTTTGAACCTGTTCCGCGACATGAGTTCAAGAAACATTCAGATGG 1167
 Db 1798 GCGGAAAAGTATATTTTGAACCTGTTCCGCGACATGAGTTCAAGAAACATTCAGATGG 1857
 QY 1168 GATGAATTAATCTTGGAGAAAGATCTGCAGAAACCTTACCGCTGACTTGTTCAGCAATG 1227
 Db 1858 GATGAATTAATCTTGGAGAAAGATCTGCAGAAACCTTACCGCTGACTTGTTCAGCAATG 1917
 QY 1228 CCTTCCCAAGAGATATGACAAATTTAAGGTAATTTGATTTCCAAACGGGT 1280
 Db 1918 CCTTCCCAAGAGATATGACAAATTTAAGGTAATTTGATTTCCAAACGGGT 1970
 RESULT 6
 ABL52519
 ID ABL52519 standard; DNA: 2307 BP.
 AC ABL52519;
 XX 17-JUL-2002 (first entry)
 DT XX
 DE C. elegans lin-61(sy223) DNA sequence SEQ ID NO:77.
 XX
 KW Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
 KM cell proliferation; nematode; cancer; mutant; gene; ds.
 XX
 OS Caenorhabditis elegans.
 XX
 PN MO200194545-A2.
 PD 13-DEC-2001.
 XX
 PE 01-JUN-2001; 2001MO-US17909.
 PR 02-JUN-2000; 2000US-208802P.
 XX
 PA (MASI) MASSACHUSETTS INST TECHNOLOGY.
 PI Horvitz HR, Davison EM, Lu X;
 XX
 DR WPI: 2002-401590/43.
 XX
 PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
 PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
 PT diagnosing and treating cell proliferative diseases such as cancer
 XX
 PS Disclosure; Page 115; 116pp; English.

XX
 CC The present invention describes a substantially pure nematode C. elegans
 CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
 CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
 CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
 CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
 CC in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
 CC cytosolic activity, and can be used in gene therapy. (I) is useful for
 CC modulating proliferation of a cell, and for identifying the compound that
 CC modulates cell proliferation. (I) can be used for diagnosing an animal
 CC (preferably, human) for the presence of the cell proliferation disease,
 CC or an increased chance of developing the disease, by measuring lin-8,
 CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
 CC animal, by measuring amount of lin-8, lin-56, lin-61 polypeptide in the sample
 CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
 CC (I) is also useful for identifying the compound that modulates cell
 CC proliferation. (I) and its mutants are useful for understanding cell
 CC proliferative diseases including cancer, as well as in diagnosing and
 CC treating cell proliferative diseases. (I) is also useful for identifying
 CC tumour suppressors in other species such as mammals and may be used to
 CC identify therapeutic compounds. The present sequence represents a
 CC C. elegans lin-61(sy223) DNA sequence from the present invention.
 XX
 SQ Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;
 Query Match 55.6%; Score 833; DB 24; Length 2307;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches: 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 448 AATTGCAATGATGCGGAATTCGTGCGCAAACTTCGCTGTCCTCAAAATTCGATGAAGGG 507
 Db 1138 AATTGCAATGATGCGGAATTCGTGCGCAAACTTCGCTGTCCTCAAAATTCGATGAAGGG 1197
 QY 508 AAGGCTCTCTAAGCAAGCATCGTTTCAAAAGTTGGACAACGCTCTTGAACATTTAATAT 567
 Db 1198 AAGGCTCTCTAAGCAAGCATCGTTTCAAAAGTTGGACAACGCTCTTGAACATTTAATAT 1257
 QY 568 TCCAAATCTACTGAAATATACGCGTATGCGCAATTCAGAAATTTGAGCAGCAATGAAT 627
 Db 1258 TCCAAATCTACTGAAATATACGCGTATGCGCAATTCAGAAATTTGAGCAGCAATGAAT 1317
 QY 628 GTATCTATCACAAAAGAAAGACTTCCGGAATCGCTTCCAGATGAGTGAAGCAGACAA 687
 Db 1318 GTATCTATCACAAAAGAAAGACTTCCGGAATCGCTTCCAGATGAGTGAAGCAGACAA 1377
 QY 688 GTCTTTAGCTCTGATCTCAATATGATATGAGAGAGGAAGCTTCTCATATTTCCGTGTT 747
 Db 1378 GTCTTTAGCTCTGATCTCAATATGATATGAGAGAGGAAGCTTCTCATATTTCCGTGTT 1437
 QY 748 GGATTTGACAGCATGATGATATCACTAAATGCGAAAAAGATATTTATGACACACA 807
 Db 1438 GGATTTGACAGCATGATGATATCACTAAATGCGAAAAAGATATTTATGACACACA 1497
 QY 808 AATTAATTTGCTCAAGCAATTAATAATGAGAAAAATCCAGATATGACTCAGACAGCTC 867
 Db 1498 AATTAATTTGCTCAAGCAATTAATAATGAGAAAAATCCAGATATGACTCAGACAGCTC 1557
 QY 868 ACATTTATCATATTAAGCAAAAAGATCAATTCATCCATGATTTGGAAAAAGTTAAGCTT 927
 Db 1558 ACATTTATCATATTAAGCAAAAAGATCAATTCATCCATGATTTGGAAAAAGTTAAGCTT 1617
 QY 928 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 987
 Db 1618 GGACAAAGTTTGAGCTCATGAGCCCTTGCTGAGCAATTCATTAACCTCCAGCTGCT 1677
 QY 988 TCGATTTCTCAAAATTTTGCAAAACCTGAAGAGATATCTTATTTGGGAATGATGTCAGAT 1047
 Db 1678 TCGATTTCTCAAAATTTTGCAAAACCTGAAGAGATATCTTATTTGGGAATGATGTCAGAT 1737
 QY 1048 GCATTTGAAGCAATTTTCTCATATCATATCAATTAATACATTTATGTTCCAGTTGTTAT 1107
 Db 1738 GCATTTGAAGCAATTTTCTCATATCATATCAATTAATACATTTATGTTCCAGTTGTTAT 1797

OY 1108 GCGGAAAAGTATTAATTTGGAACCTGTTCCCGCAGATGATCAAGAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATTAATTTGGAACCTGTTCCCGCAGATGATCAAGAACATTCAGATGG 1857
OY 1168 GATGAATACCTTGAGAAAGATCTGCAGAACCCACGCGTTGACTTCTTCAACCATG 1227
Db 1858 GATGAATACCTTGAGAAAGATCTGCAGAACCCACGCGTTGACTTCTTCAACCATG 1917
OY 1228 CCTTCCCAAGAGAGATTGACAAATTTAAGGTAATTCGATTTCCCAACGGGT 1380
Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGGTAATTCGATTTCCCAACGGGT 1970

RESULT 7

ABL52520
ID ABL52520 standard; DNA; 2307 BP.

AC ABL52520;

DT 17-JUN-2002 (first entry)

DE C. elegans lin-61(n3635) DNA sequence SEQ ID NO:78.

KW Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;

KW cell proliferation; nematode; cancer; mutant; gene; ds.

OS Caenorhabditis elegans.

PN WO200194545-A2.

PD 13-DEC-2001.

PF 01-JUN-2001; 2001WO-US17909.

PR 02-JUN-2000; 2000US-208802P.

XX (MASI) MASSACHUSETTS INST TECHNOLOGY.

PI Horvitz HR, Davison EM, Lu X;

DR WPI; 2002-401590/43.

PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
PS Disclosure: Page 116; 116pp; English.

CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (1) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (1) has
CC cytosolic activity and can be used in gene therapy. (1) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (1) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC (1) is also useful for identifying the compound that modulates cell
CC proliferation. (1) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (1) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(n3635) DNA sequence from the present invention.

SEQ Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;

Query Match 55.6%; Score 833; DB 24; Length 2307;

Best Local Similarity 100.0%; Pred. No. 0; Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 448 AATGCAATGATGCGCAAAATCGTGCCCAAACTTCGCTGCTCCAAAATTCGATGAAGG 507
Db 1138 AATGCAATGATGCGCAAAATCGTGCCCAAACTTCGCTGCTCCAAAATTCGATGAAGG 1197
OY 508 AAGCCTCTCCTAAGCAAGCATCGTTTCAAAAGTTGGACAAGCTCTTGAACATTTAATAT 567
Db 1198 AAGCCTCTCCTAAGCAAGCATCGTTTCAAAAGTTGGACAAGCTCTTGAACATTTAATAT 1257
OY 568 TCCAAATCTACTGAATAAGCGGTAGCGGAATTCAGAAATATGTGGACGCAATGAT 627
Db 1258 TCCAAATCTACTGAATAAGCGGTAGCGGAATTCAGAAATATGTGGACGCAATGAT 1317
OY 628 GTATCTATCACAAAGAAAGACTTTCGCCAATCGCTCCAGATGACAGACAGACAA 687
Db 1318 GTATCTATCACAAAGAAAGACTTTCGCCAATCGCTCCAGATGACAGACAGACAA 1377
OY 688 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAAGACCTTTCATATTCTCTGT 747
Db 1378 GTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAAGACCTTTCATATTCTCTGT 1437
OY 748 GGATTTGACGACGATCAATGATATCAACTAATGCGAAAAGAAATATATTGACACACA 807
Db 1438 GGATTTGACGACGATCAATGATATCAACTAATGCGAAAAGAAATATATTGACACACA 1497
OY 808 AATAAAATTTGCTCAAGCAATATAAAATGAGAAAATCCAGATATGACTCAGACAGCTC 867
Db 1498 AATAAAATTTGCTCAAGCAATATAAAATGAGAAAATCCAGATATGACTCAGACAGCTC 1557
OY 868 ACATTTGATCAATTAAGCAAAAGATCCATTTATCCCATGATTTGGAGAAAAGTTAAGTT 927
Db 1558 ACATTTGATCAATTAAGCAAAAGATCCATTTATCCCATGATTTGGAGAAAAGTTAAGTT 1617
OY 928 GGACAAAAGTTTGAGCTGATGACCCCTTGCTGACAAATTCATTAACCTCAGCTCGCT 987
Db 1618 GGACAAAAGTTTGAGCTGATGACCCCTTGCTGACAAATTCATTAACCTCAGCTCGCT 1677
OY 988 TCGATTCGCAAAATTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGATCCAGAT 1047
Db 1678 TCGATTCGCAAAATTTGCAAAAGTGAAGATATCTTATTTGGGAATGATGATCCAGAT 1737
OY 1048 GCACCTTGAAAGACGATTTCTCATATCATATACATTAATTAATTAATTAATTAATTAAT 1107
Db 1738 GCACCTTGAAAGACGATTTCTCATATCATATTAATTAATTAATTAATTAATTAATTAAT 1797
OY 1108 GCGGAAAAGTATTAATTTGGAACCTGTTCCCGCAGATGATCAAGAACATTCAGATGG 1167
Db 1798 GCGGAAAAGTATTAATTTGGAACCTGTTCCCGCAGATGATCAAGAACATTCAGATGG 1857
OY 1168 GATGAATACCTTGAGAAAGATCTGCAGAACCCACGCGTTGACTTCTTCAACCATG 1227
Db 1858 GATGAATACCTTGAGAAAGATCTGCAGAACCCACGCGTTGACTTCTTCAACCATG 1917
OY 1228 CCTTCCCAAGAGAGATTGACAAATTTAAGGTAATTCGATTTCCCAACGGGT 1280
Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGGTAATTCGATTTCCCAACGGGT 1970

RESULT 8

AAH04590
ID AAH04590 standard; cDNA; 872 BP.

XX AAH04590;

AC 26-JUN-2001 (first entry)

DT Human cDNA clone (5'-primer) SEQ ID NO:1425.

DE Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

XX Human; primer: detection; diagnosis; antisense therapy; gene therapy; ss.

OS Homo sapiens.

KM disease; cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000662.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-496933/54.
XX
PT New spatially-addressable set of single exon nucleic acid probes,
PT useful for measuring gene expression in sample derived from human
PT breast, comprises number of single exon nucleic acid probes -
XX
XX
XX Claim 4; SEQ ID NO 7736; 327pp + sequence listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting
XX the probes with a collection of detectably labelled nucleic acids
XX derived from mRNA of human breast, and then measuring the label
XX bound to each probe of the microarray. The probes are useful for
XX verifying the expression of regions of genomic DNA predicted to
XX encode proteins. They are useful for gene discovery, and for
XX determining predisposition and/or prognosing breast disease. Gene
XX expression analysis is useful for assessing the toxicity of chemical
XX agents on cells. The microarray of this invention presents a far greater
XX diversity of probes for measuring gene expression, with far less bias
XX than expressed sequence tag microarrays. The method is suitable for
XX rapid production of functional information from genomic sequence. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences.
SQ
Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
Query Match 1.3%; Score 20; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAAC T 1406
DB 156 GATGAGAAATTGATGAAC T 137
|||||
RESULT 11
AAI21886/C
ID AAI21886 standard; DNA; 280 BP.
XX
XX AAI21886;
XX
XX 12-OCT-2001 (first entry)
XX
XX Probe #11819 for gene expression analysis in human cervical cell sample.
XX
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer; ss.
XX

OS Homo sapiens.
XX
XX WO200157278-A2.
XX
PN
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
XX
XX
XX Claim 25; SEQ ID NO 11819; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression, the probes are therefore useful in grading and/or staging
XX of diseases of the cervix, notably cervical cancer.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
SQ
Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
Query Match 1.3%; Score 20; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAAC T 1406
DB 156 GATGAGAAATTGATGAAC T 137
|||||
RESULT 12
AAI07575/C
ID AAI07575 standard; DNA; 280 BP.
XX
XX AAI07575;
XX
XX 09-OCT-2001 (first entry)
XX
XX Probe #7566 used to measure gene expression in human breast sample.
XX
XX Probe; human; breast disease; breast cancer; development disorder; ss;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX
XX Homo sapiens.
XX
XX WO200157270-A2.
XX
XX 09-AUG-2001.
XX
XX 29-JAN-2001; 2001WO-US000661.
XX
XX 04-FEB-2000; 2000US-0180312.
XX 26-MAY-2000; 2000US-0207456.
XX 30-JUN-2000; 2000US-0608408.
XX

PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-476286/51.
XX
XX Novel single exon nucleic acid probe used to measuring gene expression
PT in a human breast .
XX
PS Claim 25; SEQ ID No 7566; 322pp; English.
XX
CC The present invention relates to novel single exon nucleic acid probes.
CC The present sequence is one such probe. The probes are useful for
CC measuring human gene expression in a human breast sample, where the probe
CC hybridizes at high stringency to a nucleic acid expressed in the human
CC breast. The probes are useful for predicting, diagnosing, grading,
CC staging, monitoring and prognosing diseases of the human breast,
CC particularly those diseases with polygenic aetiology. The diseases
CC include: breast cancer, disorders of development, inflammatory diseases
CC of the breast, fibrocystic changes, proliferative breast disease and
CC non-carcinoma tumours.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
Query Match 1.3%; Score 20; DB 22; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAACT 1406
Db 156 GATGAGAAATTGATGAACT 137
XXXXXXXXXXXXXXXXXXXX
RESULT 13
ABS15121/c
ID ABS15121 standard; DNA: 280 BP.
XX
AC ABS15121:
XX
DT 19-AUG-2002 (first entry)
XX
DE Human genome-derived single exon probe ORF from lung SEQ ID No 15112.
XX
XX Human; ds: single exon probe; asthma; lung cancer; COPD; ILD;
KW chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease; open reading frame; ORF.
XX
XX Homo sapiens.
OS
PN MO200186003-A2.
PD 15-NOV-2001.
XX
XX 30-JAN-2001; 2001WO-US00665.
XX
XX 04-FEB-2000; 2000US-180312P.
PR 26-MAY-2000; 2000US-207456P.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.
PR 27-SEP-2000; 2000US-236359P.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2002-114183/15.
XX
XX Spatially-addressable set of single exon nucleic acid probes, used to
PT measure gene expression in human lung samples .
XX
PS Claim 4; SEQ ID No 15112; 634pp; English.
XX
CC The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human lung comprising single exon nucleic acid probes having one of
CC 12614 nucleic acid sequences mentioned in the specification, or their
CC complements or the 12387 open reading frames derived from the 12614
CC probes. Also included are a microarray comprising the novel set of
CC probes: the novel set of probes which hybridize at high stringency to a
CC nucleic acid expressed in the human lung; measuring gene expression in a
CC sample derived from human lung, comprising (a) contacting the array with
CC a collection of detectably labeled nucleic acids derived from human lung
CC mRNA, and (b) measuring the label detectably bound to each probe of
CC the array; identifying exons in a eukaryotic genome, comprising
CC (a) algorithmically predicting at least one exon from genomic sequences
CC of the eukaryote; and (b) detecting specific hybridisation of detectably
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
CC having a fragment identical to the predicted exon, the probe is included
CC in the above mentioned microarray; assigning exons to a single gene,
CC comprising (a) identifying exons from genomic sequence by the method
CC above and (b) measuring the expression of each of the exons in several
CC tissues and/or cell types using hybridisation to a single exon
CC microarrays having a probe with the exon, where a common pattern of
CC expression of the exons in the tissues and/or cell types indicates that
CC the exons should be assigned to a single gene; a peptide comprising one
CC of 12011 sequences, mentioned in the specification, or encoded by the
CC probes/open reading frames (ORF). The probes are used for gene
CC expression analysis, and for identifying exons in a gene, particularly
CC using human lung derived mRNA and for the study of lung diseases
CC such as asthma, lung cancer, chronic obstructive pulmonary disease
CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic
CC and hyaline membrane disease. The present sequence is a single exon
CC probe open reading frame of the invention.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 280 BP; 78 A; 53 C; 43 G; 106 T; 0 other;
Query Match 1.3%; Score 20; DB 24; Length 280;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1387 GATGAGAAATTGATGAACT 1406
Db 156 GATGAGAAATTGATGAACT 137
XXXXXXXXXXXXXXXXXXXX
RESULT 14
ABA43943/c
ID ABA43943 standard; DNA: 425 BP.
XX
AC ABA43943:
XX

DT 01-FEB-2002 (first entry)
 XX Human breast cell single exon nucleic acid probe #2638.
 DE Human: microarray; single exon probe; gene expression; breast;
 KW disease; cancer; ss.
 OS Homo sapiens.
 XX WO200157271-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00662.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-496933/54.
 DR New spatially-addressable set of single exon nucleic acid probes,
 XX useful for measuring gene expression in sample derived from human
 PT breast, comprises number of single exon nucleic acid probes -
 XX
 PS Claim 1; SEQ ID NO 2638; 327bp + sequence listing; English.
 XX
 CC The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human breast and BT 474 cells. The method involves contacting
 CC the probes with a collection of detectably labelled nucleic acids
 CC derived from mRNA of human breast, and then measuring the label
 CC bound to each probe of the microarray. The probes are useful for
 CC verifying the expression of regions of genomic DNA predicted to
 CC encode proteins. They are useful for gene discovery, and for
 CC determining predisposition and/or prognosing breast disease. Gene
 CC expression analysis is useful for assessing the toxicity of chemical
 CC agents on cells. The microarray of this invention presents a far greater
 CC diversity of probes for measuring gene expression, with far less bias
 CC than expressed sequence tag microarrays. The method is suitable for
 CC rapid production of functional information from genomic sequence. The
 CC present sequence is a single exon nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 425 BP; 122 A; 88 C; 66 G; 149 T; 0 other;
 QY
 Db 1387 GATGAGAAATTGATGAAC 1406
 ||||||||||||||||||
 391 GATGAGAAATTGATGAAC 372

Query Match 1.3%; Score 20; DB 22; Length 425;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15
 AA112703/c
 ID AA112703 standard; DNA: 425 BP.
 XX
 AC AA112703;
 XX
 DT 12-OCT-2001 (first entry)
 XX

DE Probe #2636 for gene expression analysis in human cervical cell sample.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer; ss.
 XX
 OS Homo sapiens.
 XX WO200157278-A2.
 PN 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US00670.
 PF 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SG, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2001-488901/53.
 DR Human genome-derived single exon nucleic acid probes useful for
 XX analyzing gene expression in human cervical epithelial cells -
 PT
 PT
 XX
 PS Claim 25; SEQ ID NO 2636; 487bp; English.
 XX
 CC The present invention relates to human single exon nucleic acid probes
 CC (SENP). The present sequence is one such probe. The SENPs are derived
 CC from human HeLa cells. The SENPs can be used to produce a single exon
 CC microarray, which can be used for measuring human gene expression in a
 CC sample derived from human cervical epithelial cells. By measuring gene
 CC expression, the probes are therefore useful in grading and/or staging
 CC of diseases of the cervix, notably cervical cancer.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 XX Sequence 425 BP; 122 A; 88 C; 66 G; 149 T; 0 other;
 QY
 Db 1387 GATGAGAAATTGATGAAC 1406
 ||||||||||||||||||
 391 GATGAGAAATTGATGAAC 372

Query Match 1.3%; Score 20; DB 22; Length 425;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: September 2, 2003, 17:58:31
 Job time : 460 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 17:43:50 ; Search time 107 Seconds
(Without alignments)
6175.239 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgtcgaattcttgaat.....cgaataagacacactatga 1497

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapept 60.0

Searched: 569978 seqs, 220691566 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

Issued_Patents_NA: *
1: /cgn2_6/prodata/2/ina/5a_COMB.seq: *
2: /cgn2_6/prodata/2/ina/5b_COMB.seq: *
3: /cgn2_6/prodata/2/ina/6a_COMB.seq: *
4: /cgn2_6/prodata/2/ina/6b_COMB.seq: *
5: /cgn2_6/prodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/prodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	20	1.3	2044	4	US-09-327-138C-38
2	20	1.3	2863	4	US-09-327-138C-2
3	20	1.3	3646	4	US-09-327-138C-36
4	20	1.3	3665	4	US-09-327-138C-35
5	20	1.3	3686	4	US-09-327-138C-3
6	20	1.3	3686	4	US-09-327-138C-37
7	20	1.3	3940	4	US-09-327-138C-1
8	19	1.3	8494	4	US-08-961-527-163
9	18	1.2	742	4	US-09-482-273-76
10	18	1.2	781	4	US-09-328-475C-321
11	18	1.2	1172	4	US-08-936-165A-8
12	17	1.1	921	4	US-08-795-475-2
13	17	1.1	1005	3	US-09-446-504-79
14	17	1.1	1005	4	US-09-712-266-79
15	17	1.1	1050	4	US-09-107-532A-2184
16	17	1.1	1189	3	US-09-497-779A-7
17	17	1.1	1251	3	US-09-355-115-1
18	17	1.1	1305	4	US-09-329-234A-6
19	17	1.1	1308	2	US-08-795-475-4
20	17	1.1	1316	4	US-09-370-838-30
21	17	1.1	1618	1	US-07-726-607C-2
22	17	1.1	1618	1	US-07-843-949A-2
23	17	1.1	1618	2	US-08-218-978-2
24	17	1.1	1782	4	US-09-220-132-158
25	17	1.1	2207	4	US-09-254-504-6
26	17	1.1	2228	1	US-07-726-607C-1
27	17	1.1	2228	1	US-07-843-949A-1

28	17	1.1	2228	2	US-08-218-978-1	Sequence 1, Appl
29	17	1.1	2389	3	US-08-961-083-55	Sequence 55, Appl
30	17	1.1	2389	4	US-09-536-784-55	Sequence 5, Appl
31	17	1.1	2451	4	US-09-468-656A-9	Sequence 9, Appl
32	17	1.1	2634	3	US-08-949-386-26	Sequence 26, Appl
33	17	1.1	2634	3	US-08-450-562-26	Sequence 26, Appl
34	17	1.1	2634	4	US-08-984-709A-26	Sequence 26, Appl
35	17	1.1	2634	4	US-08-450-272-26	Sequence 26, Appl
36	17	1.1	2708	1	US-08-384-556A-10	Sequence 10, Appl
37	17	1.1	2708	1	US-08-331-355A-25	Sequence 25, Appl
38	17	1.1	2708	5	PCT-US94-12364-25	Sequence 25, Appl
39	17	1.1	2708	5	PCT-US95-07753-10	Sequence 10, Appl
40	17	1.1	2712	3	US-08-949-386-38	Sequence 38, Appl
41	17	1.1	2712	3	US-08-450-562-38	Sequence 38, Appl
42	17	1.1	2712	4	US-08-984-709A-38	Sequence 38, Appl
43	17	1.1	2712	4	US-08-450-272-38	Sequence 38, Appl
44	17	1.1	2715	3	US-08-777-147-2	Sequence 2, Appl
45	17	1.1	2970	3	US-08-949-386-37	Sequence 37, Appl

ALIGNMENTS

RESULT 1
US-09-327-138C-38
; Sequence 38, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOCs)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 2044
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: Promoter
; LOCATION: (-2759)...(-716)
US-09-327-138C-38

Query Match 1.38; Score 20; DB 4; Length 2044;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 504 AGGGAAGGCTCTCCTAAGCA 523
Db 18 AGGGAAGGCTCTCCTAAGCA 37

RESULT 2
US-09-327-138C-2
; Sequence 2, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOCs)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 2863

;; TYPE: DNA
;; ORGANISM: Mus musculus
;; FEATURE:
;; NAME/KEY: promoter
;; LOCATION: (-2759)...(104)
US-09-327-138C-2

Query Match 1.3%; Score 20; DB 4; Length 2863;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCTTAAGCA 523
Db 18 AGGGAAGGCTCTCTTAAGCA 37

RESULT 3
US-09-327-138C-36
; Sequence 36, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC5)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 3646
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotides
US-09-327-138C-36

Query Match 1.3%; Score 20; DB 4; Length 3646;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCTTAAGCA 523
Db 18 AGGGAAGGCTCTCTTAAGCA 37

RESULT 4
US-09-327-138C-35
; Sequence 35, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC5)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 35
; LENGTH: 3665
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)

;; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotides
;; OTHER INFORMATION: -80 to -60 deleted
US-09-327-138C-35

Query Match 1.3%; Score 20; DB 4; Length 3665;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCTTAAGCA 523
Db 18 AGGGAAGGCTCTCTTAAGCA 37

RESULT 5
US-09-327-138C-3
; Sequence 3, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC5)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: promoter
; LOCATION: (-2759)...(927)
US-09-327-138C-3

Query Match 1.3%; Score 20; DB 4; Length 3686;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 504 AGGGAAGGCTCTCTTAAGCA 523
Db 18 AGGGAAGGCTCTCTTAAGCA 37

RESULT 6
US-09-327-138C-37
; Sequence 37, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOC5)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; FILE REFERENCE: P07 42591 (18810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 3686
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: mutation
; LOCATION: (0)...(0)
; OTHER INFORMATION: promoter sequence -2759 to +927 with nucleotide at
; OTHER INFORMATION: position -74 mutated from T to A, nucleotide at
; OTHER INFORMATION: position -71 mutated from C to G, nucleotide at
; OTHER INFORMATION: position -69 mutated from G to C, and nucleotide
; OTHER INFORMATION: at position -66 mutated from A to T

US-09-327-138C-37

Query Match 1.3%; Score 20; DB 4; Length 3686;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 504 AGGGAAGGCTCTCCTTAAGCA 523
|||||
Db 18 AGGGAAGGCTCTCCTTAAGCA 37

RESULT 7

US-09-327-138C-1
; Sequence 1, Application US/09327138C
; Patent No. 6541244
; GENERAL INFORMATION:
; APPLICANT: AUBERNHAMMER, CHRISTOPH J.
; APPLICANT: MELMED, SHLOMO
; TITLE OF INVENTION: SUPPRESSOR OF CYTOKINE SIGNALING
; TITLE OF INVENTION: (SOCs)-3 PROMOTER AND METHODS FOR ITS USE IN GENETIC THERAPY
; TITLE OF INVENTION: IN HUMANS
; FILE REFERENCE: P07 42591 (19810-803
; CURRENT APPLICATION NUMBER: US/09/327,138C
; CURRENT FILING DATE: 1999-06-07
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3940

TYPE: DNA
ORGANISM: Mus musculus

FEATURE:
LOCATION: (-2907)...(1033)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

FEATURE:
LOCATION: (0)...(0)

CITY: Rockville

STATE: Maryland

COUNTRY: USA

ZIP: 20850

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.50 inch, 1.44MB storage

COMPUTER: HP Vectra 486/33

OPERATING SYSTEM: MSDOS version 6.2

SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/961,527

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Brookes, A. Anders

REGISTRATION NUMBER: 36,373

REFERENCE/DOCKET NUMBER: PB340P1

TELEPHONE: (301) 309-8504

TELEFAX: (301) 309-8512

INFORMATION FOR SEQ ID NO: 163:

SEQUENCE CHARACTERISTICS:

LENGTH: 8494 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

US-08-961-527-163

Query Match 1.3%; Score 19; DB 4; Length 8494;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 528 TCGTTTCAAGTTGACAA 546
|||||
Db 901 TCGTTTCAAGTTGACAA 919

RESULT 9

US-09-482-273-76/c
; Sequence 76, Application US/09482273
; Patent No. 6534631
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 71 Human Secreted Proteins
; FILE REFERENCE: P2030P1
; CURRENT APPLICATION NUMBER: US/09/482,273
; EARLIER FILING DATE: 2000-01-13
; EARLIER APPLICATION NUMBER: PCT/US99/15849
; EARLIER FILING DATE: 1999-07-14
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,922
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/092,956
; EARLIER FILING DATE: 1998-07-15
; NUMBER OF SEQ ID NOS: 267
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 76
; LENGTH: 742

TYPE: DNA
ORGANISM: Homo sapiens

FEATURE:
LOCATION: (707)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
LOCATION: (724)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
LOCATION: (724)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
LOCATION: (724)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
LOCATION: (724)

OTHER INFORMATION: n equals a,t,g, or c

FEATURE:
LOCATION: (724)

NAME/KEY: SITE
LOCATION: (726)
OTHER INFORMATION: n equals a,t,g, or c
US-09-482-273-76

Query Match 1.2%; Score 18; DB 4; Length 742;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 94 GAGAAAGAAAACCTTCT 111
|||||
Db 84 GAGAAAGAAAACCTTCT 67

RESULT 10
US-09-328-475C-321
Sequence 321, Application US/09328475C
Patent No. 6476207
GENERAL INFORMATION:
APPLICANT: Zhang, Jimmy
APPLICANT: Astel, Jon H.
APPLICANT: Carroll III, Eddie
APPLICANT: Endege, Wilson O.
APPLICANT: Ford, Donna M.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
APPLICANT: Steimann, Kathleen E.
TITLE OF INVENTION: GENES AND GENE EXPRESSION PRODUCTS THAT
TITLE OF INVENTION: ARE DIFFERENTIALLY REGULATED IN PROSTATE CANCER
FILE REFERENCE: 1532.002/200130.463
CURRENT APPLICATION NUMBER: US/09/328,475C
CURRENT FILING DATE: 1999-06-09
NUMBER OF SEQ ID NOS: 341
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 321
LENGTH: 781
TYPE: DNA
ORGANISM: Homo Sapien
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(781)
OTHER INFORMATION: n = A,T,C or G
US-09-328-475C-321

Query Match 1.2%; Score 18; DB 4; Length 781;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 96 GAAAGGAAAACCTTCTT 113
|||||
Db 551 GAAAGGAAAACCTTCTT 568

RESULT 11
US-08-936-165A-8
Sequence 8, Application US/08936165A
Patent No. 6348582
GENERAL INFORMATION:
APPLICANT: Black, Michael
APPLICANT: Burnham, Martin
APPLICANT: Hodgson, John
APPLICANT: Knowles, David
APPLICANT: Lonetto, Michael
APPLICANT: Nicholas, Richard
APPLICANT: Pratt, Julie
APPLICANT: Reichard, Richard
APPLICANT: Rosenberg, Martin
APPLICANT: Ward, Judith
TITLE OF INVENTION: NO. 6348582zel Prokaryotic Polynucleotides,
TITLE OF INVENTION: Polypeptides and Their Uses
NUMBER OF SEQUENCES: 534
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SmithKline Beecham Corporation

SMREET: 709 Swedeland Road
CITY: King of Prussia
STATE: PA
COUNTRY: USA
ZIP: 19406-0939
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/936,165A
FILING DATE: 24-SEP-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/027,032
FILING DATE: 24-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Gimm1, Edward R
REGISTRATION NUMBER: 38,891
REFERENCE/DOCKET NUMBER: P50549
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-4478
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 1172 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-936-165A-8

Query Match 1.2%; Score 18; DB 4; Length 1172;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CTTGAAGCTGCTGACATG 1308
|||||
Db 810 CTTGAAGCTGCTGACATG 827

RESULT 12
US-08-795-475-2
Sequence 2, Application US/08795475
Patent No. 5965390
GENERAL INFORMATION:
APPLICANT: Bjvrck, Lars
APPLICANT: Sjvrck, Ulf
TITLE OF INVENTION: PROTEIN L AND HYBRID PROTEINS THEREOF
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESSES:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/795,475
FILING DATE: 11-FEB-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 100084.402D1
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 921 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ORIGINAL SOURCE:
US-08-795-475-2
Escherichia coli LE392/PHDL, DSM 7054

Query Match 1.1%; Score 17; DB 2; Length 921;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1144 GAGTCAAGACATT 1160
|||||
Db 757 GAGTCAAGACATT 773

RESULT 13
US-09-446-504-79
Sequence 79, Application US/09446504
Patent No. 6218150
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazue
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo
APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/446, 504
CURRENT FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1005
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-446-504-79

Query Match 1.1%; Score 17; DB 3; Length 1005;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 309 TTGGCTCATATTTAT 325
|||||
Db 5 TTGGCTCATATTTAT 21

RESULT 14
US-09-712-266-79
Sequence 79, Application US/09712266
Patent No. 6333158
GENERAL INFORMATION:
APPLICANT: UEMORI, Takashi
APPLICANT: SATO, Yoshimi
APPLICANT: FUJITA, Tomoko
APPLICANT: MIYAKE, Kazue
APPLICANT: MUKAI, Hiroyuki
APPLICANT: ASADA, Kiyozo

APPLICANT: KATO, Ikunoshin
TITLE OF INVENTION: DNA POLYMERASE-RELATED FACTORS
FILE REFERENCE: 1422-408PCT
CURRENT APPLICATION NUMBER: US/09/712, 266
CURRENT FILING DATE: 2000-11-15
PRIOR APPLICATION NUMBER: US 09/446, 504
PRIOR FILING DATE: 1999-12-23
PRIOR APPLICATION NUMBER: PCT/JP98/02845
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: JP 9-187496
PRIOR FILING DATE: 1997-06-26
PRIOR APPLICATION NUMBER: JP 9-320692
PRIOR FILING DATE: 1997-11-27
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 1005
TYPE: DNA
ORGANISM: Pyrococcus furiosus
US-09-712-266-79

Query Match 1.1%; Score 17; DB 4; Length 1005;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 309 TTGGCTCATATTTAT 325
|||||
Db 5 TTGGCTCATATTTAT 21

RESULT 15
US-09-107-532A-2184
Sequence 2184, Application US/09107532A
Patent No. 6583275
GENERAL INFORMATION:
APPLICANT: Lynn A Doucette-Stamm and David Bush
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
NUMBER OF SEQUENCES: 7310
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107, 532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085, 598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Arinello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GRC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2184:
SEQUENCE CHARACTERISTICS:
LENGTH: 1050 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)

```

;
;   HYPOTHETICAL: NO
;   ANTI-SENSE: NO
;   ORIGINAL SOURCE:
;   ORGANISM: Enterococcus faecium
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (8) LOCATION 1...1050
;   SEQUENCE DESCRIPTION: SEQ ID NO: 2184:
US-09-107-532A-2184

```

```

Query Match      1.1%; Score 17; DB 4; Length 1050;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1158 ATTCAGATGGAGTGAAT 1174
          |||||
Db      947 ATTCAGATGGAGTGAAT 963

```

Search completed: September 2, 2003, 20:33:02
 Job time : 110 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 13:43:09 ; Search time 455 Seconds
(without alignments)
8881.448 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497
Sequence: 1 atgctgaattctgaaat.....cgaaaagracactatga 1497

Scoring table: IDENTITY_NIC
Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.Geneseq_19Jun03:*

- 1: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSL/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1497	100.0	1497	24	ABL52497
2	1495.4	99.9	1497	24	ABL52515
3	1495.4	99.9	1497	24	ABL52516
4	1495.4	99.9	1497	24	ABL52517
5	1103	73.7	2307	24	ABL52518
6	1103	73.7	2307	24	ABL52519
7	1103	73.7	2307	24	ABL52520
8	58.8	3.9	2969	22	AA513628

9	57.4	3.8	3590	21	AA294124
10	55.8	3.7	2564	25	ABX08841
11	55.2	3.7	2403	22	AA527361
12	55.2	3.7	2968	23	ABV20957
13	55.2	3.7	2968	23	ABV26806
14	55.2	3.7	3280	24	ABQ54484
15	53.4	3.6	1035	22	AA563158
16	53.4	3.6	1838	23	AA568094
17	53.4	3.6	2081	23	AA568090
18	53.4	3.6	2855	18	AAV00011
19	53.4	3.6	2855	18	AAV97297
20	53.4	3.6	2855	20	AAV61313
21	53.4	3.6	3255	18	AAV97299
22	53.4	3.6	3255	18	AAV00013
23	53.4	3.6	3255	18	AAV61315
24	53.4	3.6	3327	18	AAV97298
25	53.4	3.6	3327	18	AAV00012
26	53.4	3.6	3327	20	AAV61314
27	52	3.5	204	22	ABA50121
28	52	3.5	204	22	ABA35090
29	52	3.5	204	22	AA122966
30	51.8	3.5	3105	22	AAH16246
31	51.6	3.4	494	22	ABA44962
32	51.6	3.4	494	22	ABA25176
33	51.6	3.4	494	22	AA113745
34	51.4	3.4	4171	23	AA592635
35	50.6	3.4	2380	22	AA160518
36	50.6	3.4	2555	22	AA559625
37	50.6	3.4	2781	22	AA158732
38	50.6	3.4	3071	24	ABA93718
39	50.6	3.4	3530	25	ACC50563
40	50.6	3.4	3530	25	ABZ71320
41	50.2	3.4	3065	18	AAV00010
42	50.2	3.4	3065	18	AAV00014
43	50.2	3.4	3065	20	AAV61316
44	48	3.2	418	24	ABL68929
45	48	3.2	4590	22	AAH24065

ALIGNMENTS

RESULT 1
ABL52497
ID ABL52497 standard; DNA: 1497 BP.
AC ABL52497;
XX 17-JUL-2002 (first entry)
DT
XX
DE Caenorhabditis elegans lin-61 nucleotide sequence SFG ID NO:6.
KW Caenorhabditis elegans; lin-6; lin-56; lin-61; tumour suppressor;
KW cell proliferation; nematode; cancer; gene; ds.
XX
OS Caenorhabditis elegans.
XX
FH Key location/Qualifiers
FT CDS 1..1497
FT /*tag= a
FT /*product= "LIN-61 protein"
XX
PN W0200194545-A2.
PD 13-DEC-2001.
XX
XX 01-JUN-2001; 2001MO-US17909.
XX 02-JUN-2000; 2000US-208802P.
XX (MAST) MASSACHUSETTS INST TECHNOLOGY.
XX Horvitz HR, Davison EM, Lu X;
PI

Haematopoietic stem
Angiogenesis-asso
CDNA encoding nove
Human prostate exp
Human prostate exp
Human ovarian anti
Human purified sec
DNA encoding novel
DNA encoding novel
Human sex comb on
Human sex comb on
Human Scm protein
Human sex comb on
Human sex comb on
Human sex comb on
Human Scm protein
Human breast cell
Probe #13516 for g
Probe #12899 for g
Human CDNA sequenc
Human breast cell
Probe #3642 for ge
Probe #3678 for ge
DNA encoding novel
Human polynucleoti
Human cell cycle a
Human polynucleoti
Human transmembran
Human secreted pro
Secreted protein-e
Mouse sex comb on
Mouse sex comb on
Mouse Scm protein
Kidney cancer rela
Yeast AOD9604-asso

XX	C. elegans mutant lin-61 DNA sequence lin-61(n3446) SEQ ID NO:73.
DE	
XX	
KW	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KV	cell proliferation; nematode; cancer; mutant; gene; ds.
XX	
OS	Caenorhabditis elegans.
XX	
EH	Key
FT	Location/Qualifiers
CDS	1..1497
	/*tag= a
FT	/product= "LIN-61(n3446) protein"
XX	
PN	MO200194545-A2.
XX	
PD	13-DEC-2001.
XX	
PD	01-JUN-2001; 2001MO-US17909.
PF	
XX	
PR	02-JUN-2000; 2000US-208802P.
XX	
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.
XX	
PI	Horvitz HR, Davison EM, Lu X;
XX	
XX	WPI: 2002-401590/43.
DR	P-PSDB; ABB78697.
XX	
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT	pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT	diagnosing and treating cell proliferative diseases such as cancer
XX	
XX	Claim 27: Page 112-113; 116pp; English.
XX	

Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;

	Query Match	99.9%	Score 1495.4;	DB 24;	Length 1497;
	Best Local Similarity	99.9%;	Pred. No. 0;		
	Matches 1496;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Oy	1 ATGTCGTGAATTTCTGAAAAATTGTCAGAGCTTACAAAAATCGCAGAAAACCTGGATTAG	60			
Dd	1 ATGTCGTGAATTTCTGAAAAATTGTCAGAGCTTACAAAAATCGCAGAAAACCTGGATTAG	60			
Oy	61 ACCTACTYTGTEGGAGATCCTATTATACATCAGATTCGAGAAAGGAAAAACCTCTTCAATTCGA	120			
Dd	61 ACCACTCTGTGGGGAATCCTATTATACATCAGATTCGAGAAAGGAAAAACCTCTTCAATTCGA	120			
Oy	121 GTTGAGAGCATTCAAATCGTAACCTTACAGTTAATTTTAAAGANTCGTGMAAGGAGATT	180			
Dd	121 GTTGAGAGCATTCAAATCGTAACCTTACAGTTAATTTTAAAGANTCGTGMAAGGAGATT	180			

QY	181	ATCTTCCAAACAGTGTCCATGATTTATGACAAAGACTGCATTTGGATTCCAAGTCAGATG	240
Db	181	ATCTTCGAACAGTGTCCATGATTTATGACAAAGACTGCATTTGGATTCCAAGTCAGATG	240
QY	241	TTTGCAGCAATTTGAAAAAGTTTGGGGAATACAGAGTCTGGGCTCAGTTTATCGAGCTGAC	300
Db	241	TTTGCAGCAATTTGAAAAAGTTTGGGGAATACAGAGTCTGGGCTCAGTTTATCGAGCTGAC	300
QY	301	ACGAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAAAACGCCGCAATG	360
Db	301	ACGAATTTTGGCTCAATATTTTATCGGACGATATGTTTGGTTTGGCAAAACGCCGCAATG	360
QY	361	AGTATGCCAATATGAGATAAAATGTATATCTCCGCGCTTGGAATCAACGAAATPAC	420
Db	361	AGTATGCCAATATGAGATAAAATGTATATCTCCGCGCTTGGAATCAACGAAATPAC	420
QY	421	CAAAATGATATGATAATTTATGTAAATTAATTCAGATTGAGCGAAATCGTCGCGCAACT	480
Db	421	CAAAATGATATGATAATTTATGTAAATTAATTCAGATTGAGCGAAATCGTCGCGCAACT	480
QY	481	TGCGTGTCTCCAAAATTCGATGAAGGGAAGGCTCTCTAGCAAGCATCGTTTCAAAGTT	540
Db	481	TGCGTGTCTCCAAAATTCGATGAAGGGAAGGCTCTCTAGCAAGCATCGTTTCAAAGTT	540
QY	541	GGACAAGCTCTGGAACCTATTAATTTATCCAACTCTACTGAAATACGGGTAGCCGCAAT	600
Db	541	GGACAAGCTCTGGAACCTATTAATTTATCCAACTCTACTGAAATACGGGTAGCCGCAAT	600
QY	601	CAAGAAATATGTGAGCAGCAGATGATATCTATCACAAAGAAAGACTTCCCGAATCG	660
Db	601	CAAGAAATATGTGAGCAGCAGATGATATCTATCACAAAGAAAGACTTCCCGAATCG	660
QY	661	CTTCAGATGAGATGACGACAGCAAGTCTTATGCTGTGATCTCAATATTTGGATGAC	720
Db	661	CTTCAGATGAGATGACGACAGCAAGTCTTATGCTGTGATCTCAATATTTGGATGAC	720
QY	721	GAGGGAACCTCTTCATATTTTCCGTGTTGGATTTGACAGAGTCAATGATATCAACTAAAT	780
Db	721	GAGGGAACCTCTTCATATTTTCCGTGTTGGATTTGACAGAGTCAATGATATTTGATGAC	780
QY	781	GCGAAAAAGGAAATATTTGAGCACACAAATTAATTTGCTCAAGCAATTAATAATTTGGAGAA	840
Db	781	GCGAAAAAGGAAATATTTGAGCACACAAATTAATTTGCTCAAGCAATTAATAATTTGGAGAA	840
QY	841	AATCCAAAGATATGACTCAGACGAGTCCATTTATCATTAATGACAAAAGATCCAAATGAT	900
Db	841	AATCCAAAGATATGACTCAGACGAGTCCATTTATCATTAATGACAAAAGATCCAAATGAT	900
QY	901	CCCATGATTTGGAGAAAAGTTAAGGTTGACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAGAAAAGTTAAGGTTGACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
QY	961	CAGCAATTCATTAACCTCCACGTCGCTCGATTCGAAATTTTTCAAAACGGAAGATAT	1020
Db	961	CAGCAATTCATTAACCTCCACGTCGCTCGATTCGAAATTTTTCAAAACGGAAGATAT	1020
QY	1021	CTTATTTGGGAATGAGATGTCACGATGACACTTGAAGCAAGTTTCTTATTCATATCAAT	1080
Db	1021	CTTATTTGGGAATGAGATGTCACGATGACACTTGAAGCAAGTTTCTTATTCATATCAAT	1080
QY	1081	AATACATTTATGTTTCCAGTTGGTTATGCGGAAAAGTATTAATTTGGAACCTTGTTCGCCCA	1140
Db	1081	AATACATTTATGTTTCCAGTTGGTTATGCGGAAAAGTATTAATTTGGAACCTTGTTCGCCCA	1140
QY	1141	GATGAGTTCAAGGAACATTCAGATGGGATTAATCTTGGAAGAAAGATCTGCAGAAAC	1200
Db	1141	GATGAGTTCAAGGAACATTCAGATGGGATTAATCTTGGAAGAAAGATCTGCAGAAAC	1200
QY	1201	CTACCGCTTGACTTGTTCAGCCCAATGCTTCCCAAGAGAGATTAAGCAATTTAAGCTA	1260
Db	1201	CTACCGCTTGACTTGTTCAGCCCAATGCTTCCCAAGAGAGATTAAGCAATTTAAGCTA	1260

QY	541	GGACAAAGCTGTGAAACTTAATTAATTCACAACTCTACTGAAATACGGGTGCGCAATT	600
Db	541	GGACAAAGCTGTGAAACTTAATTAATTCACAACTCTACTGAAATACGGGTGCGCAATT	600
QY	601	CAAGAAATATGTGAGCAGCAGAAATGATATCTATCACAAGAAAGACTTCCGAATCG	660
Db	601	CAAGAAATATGTGAGCAGCAGAAATGATATCTATCACAAGAAAGACTTCCGAATCG	660
QY	661	CTTCCAGATGAGATGACGACAGACAGCTCTTACGCTCGATGCTCAATATTTGGATAGAC	720
Db	661	CTTCCAGATGAGATGACGACAGACAGCTCTTACGCTCGATGCTCAATATTTGGATAGAC	720
QY	721	GAGGGAAGCTCTTCATATTTCCGTTGGATTTCAGCAGCTCAATGATATCAACTAAAT	780
Db	721	GAGGGAAGCTCTTCATATTTCCGTTGGATTTCAGCAGCTCAATGATATCAACTAAAT	780
QY	781	GGCAAAAAAGGAATATATTGAGCACACACAATTAATTGTCAAGCAATTAATAATGAGAA	840
Db	781	GGCAAAAAAGGAATATATTGAGCACACACAATTAATTGTCAAGCAATTAATAATGAGAA	840
QY	841	AATCCAAGATATGACTCAGACGAGCTCACATTTGATCAATTAGCAAAAGATCCAAATTGAT	900
Db	841	AATCCAAGATATGACTCAGACGAGCTCACATTTGATCAATTAGCAAAAGATCCAAATTGAT	900
QY	901	CCCATGATTTGGAAAAAGTTAAGGTTGGACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
Db	901	CCCATGATTTGGAAAAAGTTAAGGTTGGACAAAAGTTTGAGCTCATCGACCCCTTGCT	960
QY	961	CAGCAATTCATTAACCTCCAGCTGCGCTCGAATTCGCAATTTTGGAAAACGGAAGGATAT	1020
Db	961	CAGCAATTCATTAACCTCCAGCTGCGCTCGAATTCGCAATTTTGGAAAACGGAAGGATAT	1020
QY	1021	CTTATTTGGGAATGAGATGCTCCAGATGACACTTGTAAGACAGTTTCCATTCATATCAAT	1080
Db	1021	CTTATTTGGGAATGAGATGCTCCAGATGACACTTGTAAGACAGTTTCCATTCATATCAAT	1080
QY	1081	AATACATTTAATGTTCCCACTGGTGTTATGCGGAAAAAGTAAATTTGGAACTTGTCCCGCA	1140
Db	1081	AATACATTTAATGTTCCCACTGGTGTTATGCGGAAAAAGTAAATTTGGAACTTGTCCCGCA	1140
QY	1141	GATGAGTTTCAAGAGAACATTCAGATGGGATGAATACTTGGAGAAAGAAATCTGCAGAAAC	1200
Db	1141	GATGAGTTTCAAGAGAACATTCAGATGGGATGAATACTTGGAGAAAGAAATCTGCAGAAAC	1200
QY	1201	CTACCGCTGTGACTGTGTTCAACCAATGCTTCCCAAGAGAAATTAGCAAAATTTAAGTA	1260
Db	1201	CTACCGCTGTGACTGTGTTCAACCAATGCTTCCCAAGAGAAATTAGCAAAATTTAAGTA	1260
QY	1261	AATTCGATTTCCAAACGGGTGAGCACTACGCTTGGAAGCTGTGACATGTGTGAANAATGAG	1320
Db	1261	AATTCGATTTCCAAACGGGTGAGCACTACGCTTGGAAGCTGTGACATGTGTGAANAATGAG	1320
QY	1321	TTTATTTGTCAGCTACAGTGAATACGTTCAATGAGAGAGCTGATTAATGTCAATTTGCAC	1380
Db	1321	TTTATTTGTCAGCTACAGTGAATACGTTCAATGAGAGAGCTGATTAATGTCAATTTGCAC	1380
QY	1381	GGCGGGGATGAAGAAATTTGATGAACCTGATGATGAGACCTCCCAATGATATCTACCGATA	1440
Db	1381	GGCGGGGATGAAGAAATTTGATGAACCTGATGATGAGACCTCCCAATGATATCTACCGATA	1440
QY	1441	GGATGCTGTGAAGCGCAGCTTATGTTCTACAACTCCGAAAAAGTACAACTATTGGA	1497
Db	1441	GGATGCTGTGAAGCGCAGCTTATGTTCTACAACTCCGAAAAAGTACAACTATTGGA	1497
RESULT 5			
ABL52518			
ID	ABL52518	standard; DNA; 2307 BP.	
XX	ABL52518;		
XX			
DT	17-JUL-2002	(first entry)	
XX			

DE	C. elegans lin-61 DNA sequence SEQ ID NO:76.
XX	
KW	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM	cell proliferation; nematode; cancer; mutant; gene; ds.
XX	
OS	Caenorhabditis elegans.
XX	
PM	M0200194545-A2.
PD	13-DEC-2001.
PF	01-JUN-2001; 2001MO-USI7909.
PR	02-JUN-2000; 2000US-208802P.
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.
XX	
PI	Horvitz HR, Davison EM, Lu X;
DR	WPI; 2002-401590/43.
XX	
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for diagnosing and treating cell proliferative diseases such as cancer -
PS	Disclosure; Page 114-115; 116pp; English.
XX	
CC	The present invention describes a substantially pure nematode C. elegans (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide (see ABB78650), where the polypeptides comprise at least 130, 110, 130 contiguous amino acids of the 386, 322 and 498 amino acid sequences given in ABB78648 to ABB78650, and modulate cell proliferation. (I) has cytostatic activity, and can be used in gene therapy. (I) is useful for modulating proliferation of a cell, and for identifying the compound that modulates cell proliferation. (I) can be used for diagnosing an animal (preferably, human) for the presence of the cell proliferation disease, or an increased chance of developing the disease, by measuring lin-8, lin-56 or lin-61 nucleic acid expression in a sample obtained from the animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample, using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample. (I) is also useful for identifying the compound that modulates cell proliferation. (I) and its mutants are useful for understanding cell proliferative diseases including cancer, as well as in diagnosing and treating cell proliferative diseases. (I) is also useful for identifying tumour suppressors in other species such as mammals and may be used to identify therapeutic compounds. The present sequence represents a C. elegans lin-61 DNA sequence from the present invention.
SQ	Sequence 2307 BP; 740 A; 393 C; 466 G; 708 T; 0 other;
	Query Match 73.7%; Score 1103; DB 24; Length 2307;
	Best Local Similarity 81.3%; Pred. No. 2,2e+283;
	Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5
Db	
OY	1 ATGCTGAATTCTCAAAAATGTGCAGACTAACAACAAATCGCAGAGAAATCGCATTAAG 60 467 ATGTGTGAATTTCTAAAAATGTGCAGACTAACACAAATAATCGCAGAGAAATCGCATTAAG 526
Db	
OY	61 ACCTACTTGTGGGAATCCATTTTACATCAGTTTCGAGAAAGAAAACCTCTTCAATCCA 120 527 ACCTACTTGTGGGAATCCATTTTACATCAGTTTCGAGAAAGAAAACCTCTTCAATCCA 586
OY	121 GTTGAGAGATTCATACGTAACCTTACATGTAATTTTAACGATGGGTGAAGAAGAGAGTT 180 587 GTTGAGAGATTCATACGTAACCTTACATGTAATTTTAACGATGGGTGAAGAAGAGAGTT 646
Db	
OY	181 ATC-----TTCGAACA 192 647 ATTCGAGATTCATATTTGTTTCGTAATCGGTTTTAAATACATTTTTTTGTAAGTTCCAACA 706
OY	193 GTGCTCATGATTTATGACAGAACTGCAGTTTCGATTCGAAGTCAGATGTTTTCGACAGAAAT 252

Db	707	GTGTCACATGATTTATGACAGAACTCGGATTCGATTCAAGTCAGATGGTTTGACGAATT	766
OY	253	GAAGAAAGTTTCGGATACAGAGTTCTGCGCTCAGTTTATCGGAGCGACACGAATTTTGG	312
Db	767	GAAGAAAGTTTCGGATACAGAGTTCTGCGCTCAGTTTATCGGAGCTGCACGAATTTTGG	826
OY	313	CTCAATATTTTATCGGAGCATATGTTGGTTTGGCAAA	350
Db	827	CTCATATTTTATCGGAGCATATGTTGGTTTGGCAAAATGAAGTTGACGCTCAGCTCTT	886
OY	351	-----	350
Db	887	TCATCTATTCTAAATTAATAATGGTTCGTTACATTAATAATCTAGAGAACATCGTATTA	946
OY	351	-----CGCGCAATGAGTAT	366
Db	947	AAACTCGAAACATTTTGATATATAGTAAATTTGACATTTACAGCGCGCAATGAGTAT	1006
OY	367	CCCAATATGATTAATAATTTGATATGCTCGCGCGCTTGCAATCAACGAAGATACCAAAAT	426
Db	1007	CCCAATATGATTAATAATTTGATATGCTCGCGCGCTTGCAATCAACGAAGATACCAAAAT	1066
OY	427	GATATGGTAATTTATGTAAAT	447
Db	1067	GATATGGTAATTTATGTAAATGTAAGTTGTTTTTTCGAAATTTAGTTAATATCATCT	1128
OY	448	-----AATTGCAATGATGGCGAAATCGTGGCGCAAACTTCGCTCTCCAAAT	496
Db	1127	CACAACCTTCAGAAATTTGCATTTATGGCGAAATCGTGGCGCAAACTTCGCTCTCCAAAT	1186
OY	497	TCGATGAAGGGAAGGCTCTCCTTAAGCAAGCATCGTTCAAGTTTGACACAGCTCTTGAC	556
Db	1187	TCGATGAAGGGAAGGCTCTCCTTAAGCAAGCATCGTTCAAGTTTGACACAGCTCTTGAC	1248
OY	557	TATTAATATTTTCAATCTCAATCTGAGAAATAGCGGTAGCGCGCAATCAAAATATGTGAC	616
Db	1247	TATTAATATTTTCAATCTCAATCTGAGAAATAGCGGTAGCGCGCAATCAAAATATGTGAC	1306
OY	617	GACGAATGAATGTATCTATCACAAAAGAAAGCTTCCGAATTCGCTTCAGATGCAAGATG	676
Db	1307	GACGAATGAATGTATCTATCACAAAAGAAAGCTTCCGAATTCGCTTCAGATGCAAGATG	1366
OY	677	ACGACAGCAAGTCTTATAGCTCTGAGTCAATATTTGGATAGACAGGGAAGCTTCCTCA	736
Db	1367	ACGACAGCAAGTCTTATAGCTCTGAGTCTCAATATTTGGATAGACAGGGAAGCTTCCTCA	1428
OY	737	TATTTCTCTGTGGATTTGCACAGTCATGATATCAACTAATTCGCAAAAGAAATTA	796
Db	1427	TATTTCTCTGTGGATTTGCACAGTCATGATATCAACTAATTCGCAAAAGAAATTA	1486
OY	797	TTGAGCACACAATAAATTTCTGAAGCAATTAATAATGGAGAAATCCAGATATGATCT	856
Db	1487	TTGAGCACACAATAAATTTCTGAAGCAATTAATAATGGAGAAATCCAGATATGATCT	1546
OY	857	CAGACGAGCTCACATTTGATCAATTTAGCAAAAGATCCATTTGATCCATGATTTTGGAGAA	916
Db	1547	CAGACGAGCTCACATTTGATCAATTTAGCAAAAGATCCATTTGATCCATGATTTTGGAGAA	1606
OY	917	AAAGTTAAGTTGGACAAAAGTTTGAGCTCATCGACCCCTTGCGTCAGCAATTCATAAC	976
Db	1607	AAAGTTAAGTTGGACAAAAGTTTGAGCTCATCGACCCCTTGCGTCAGCAATTCATAAC	1666
OY	977	TCGACGTCGCTCGATTTCTCAAAATTTTGGCAAAACGAAGATATCTTTATTTGGGAATGG	1038
Db	1667	TCGACGTCGCTCGATTTCTCAAAATTTTGGCAAAACGAAGATATCTTTATTTGGGAATGG	1728
OY	1037	ATGTCACGATGCACTTGAAGACAGTTTCTATATTCATATCAATTAATCATTTATGTTCC	1098
Db	1727	ATGTCACGATGCACTTGAAGACAGTTTCTATATTCATATCAATTAATCATTTATGTTCC	1788
OY	1097	CAGTTGGTTATGCGGAAAATATATATTTGGAACCTTGTCCGCGACAGATGATTCAAAGAA	1156
Db	1787	CAGTTGGTTATGCGGAAAATATATATTTGGAACCTTGTCCGCGCGAGATGATTCAAAGAA	1848

OY	1157	CATTGAGATGGGATGAACTACTTTGGAGAAAGAAATCTGCAGAAACCCCTACCGCTTGACCTGTG	1216
Db	1847	CATTGAGATGGGATGAACTACTTTGGAGAAAGAAATCTGCAGAAACCCCTACCGCTTGACCTGTG	1906
OY	1217	TCAAGCCCAATGCCCTTCCCAAGAGAGATTAGCAAAATTAAAGGTAATTCTGATTTCCAAAC	1276
Db	1907	TCAAGCCCAATGCCCTTCCCAAGAGAGATTAGCAAAATTTAAAGGTAATTCTGATTTCCAAAC	1966
OY	1277	G-----	1277
Db	1967	GGGTGTTTATTCGTTTGGAGATTGTTTTCACATATTAATGTTATTCATATTTGTTCTT	2026
OY	1278	-----GGTAGACACTACGCCCTTGAAGCTGCTGACATGCTGAGAAATTCAGTTATTTGTC	1330
Db	2027	GTTTAAAGGTAGGAGCTACGCCCTTGAAGCTGCTGACATGCTGAGAAATTCAGTTATTTGTC	2086
OY	1331	CAGCTACGATGAATAGCTTCATGGAAGACATGATTAATGCAATTTGGAGCGCTGGGANTG	1390
Db	2087	CAGCTACGATGAATAGCTTCATGGAAGACATGATTAATGCAATTTGGAGCGCTGGGANTG	2146
OY	1391	AAGAATTTTGATGAACACTGTATGATGTGA-----	1418
Db	2147	AAGAATTTTGATGAACACTGTATGATGTGAAGTTTATCATGACCGAAGACATTTTTC	2206
OY	1419	-----CTCCCATGATATTTCTACCCATATGATGATGCTGTGAACGCC	1456
Db	2207	AATGAAAATTCATCATTTTCAGCTCCCAATTCATCAACATGATGATGTGAACGCC	2266
OY	1457	ACAGTTATGTTCTACCAACCTCCGAAAAAGTACCACTATTGA	1497
Db	2267	ACAGTTATGTTCTACCAACCTCCGAAAAAGTACCACTATTGA	2307
RESULT 6			
ABL52519			
ID	ABL52519	standard; DNA; 2307 BP.	
XX	ABL52519;		
AC	17-JUL-2002	(first entry)	
XX			
DT			
XX			
XX			
DE	C. elegans lin-61(sy223)	DNA sequence SEQ ID NO:77.	
XX			
KM	Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;		
XX	cell proliferation; nematode; cancer; mutant; gene; ds.		
OS	Caenorhabditis elegans.		
XX			
PN	WO200194545-A2.		
XX			
XX	13-DEC-2001.		
XX			
PF	01-JUN-2001; 2001WO-US17909.		
XX			
PR	02-JUN-2000; 2000US-208802P.		
XX			
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.		
XX			
PI	Horvitz HR, Davison EM, Lu X;		
XX			
DR	WPI: 2002-401590/43.		
XX			
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor		
PT	pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for		
PT	diagnosis and treating cell proliferative diseases such as cancer		
XX			
PS	Disclosure: Page 115; 116pp; English.		
XX			
XX	The present invention describes a substantially pure nematode C. elegans		
CC	(Caenorhabditis elegans) nucleic acid (1) encoding a LIN-8 polypeptide		
CC	(see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide		
CC	(see ABB78650), where the polypeptides comprise at least 130, 110, 130		

CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB78648 to ABB78650, and modulate cell proliferation. (1) has
CC cytosolic activity, and can be used in gene therapy. (1) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (1) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring lin-8, lin-56 or lin-61 polypeptide in the sample
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (1) is also useful for identifying the compound that modulates cell
CC proliferation. (1) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (1) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(sy223) DNA sequence from the present invention.
XX

Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other:

Query Match 73.7%; Score 1103; DB 24; Length 2307;

Best Local Similarity 81.3%; Pred. No. 2.2e-283;

Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

```
1  ATGCTGGAATTTCTGAAAATTTCTCAGAGCTACAAAAATCGACAGAAAACCTCGATAG  60
   |||||||
467  ATGCTGGAATTTCTGAAAATTTCTCAGAGCTACAAAAATCGACAGAAAACCTCGATAG  526
   |||||||
61  ACCCTCTGTGGGAATCCATTTACATCAGTTCGAGAAAAGAAAACCTCTTCATTTCA  120
   |||||||
527  ACCCTCTGTGGGAATCCATTTACATCAGTTCGAGAAAAGAAAACCTCTTCATTTCA  586
   |||||||
121  GTTGAAGCATTCATCGTAACCTTACAGTTAAATTTAAAGAAATGCGTGAAGAGAGCTT  180
   |||||||
587  GTTGAAGCATTCATCGTAACCTTACAGTTAAATTTAAAGAAATGCGTGAAGAGAGCTT  646
   |||||||
181  ATC-----TTGCAACA  192
   |||
647  ATCGTAGACTCAATTTGTTCTGTAATTCGTTTTAAATAACAATTTTGTAGTTCAAAA  706
   |||
193  GTGTGCATGATTTATGACAAAGACGATTCGATTCAGATTCAGATTCGATTCAGATTC  252
   |||||||
707  GTGTGCATGATTTATGACAAAGACGATTCGATTCAGATTCAGATTCGATTCAGATTC  766
   |||||||
253  GAAAAAGTTTGGGATACAGAGTTCTGGCTCAGTTATTCGAGCTGACAGAAATTTTGG  312
   |||||||
767  GAAAAAGTTTGGGATACAGAGTTCTGGCTCAGTTATTCGAGCTGACAGAAATTTTGG  826
   |||||||
313  CTCAATATTTTTCGAGCATATGTTTGGTTGGCAAA-----  350
   |||||||
827  CTCAATATTTTTCGAGCATATGTTTGGTTGGCAAAAGTAAGTTGACGCTCAGCTCTT  886
   |||||||
351  -----  350
   |||
887  TCTACTATTTCTAAATTAATATGTTCTGTACATTAATTTCTAGAGAAATCGATTTA  946
   |||
351  -----GCGCCGATGAGTAT  366
   |||||||
947  AAACCTCGAAACATTTGTATTAATAGTAAATTTGAACATTTTCAGCGCCCAATGAGTAT  1006
   |||||||
367  CCCAATATGATTAATTTGTATATGCTCGCGGCTTGCATCAACGAAATATCCAAAAT  426
   |||||||
1007  CCCAATATGATTAATTTGTATATGCTCGCGGCTTGCATCAACGAAATATCCAAAAT  1066
   |||||||
427  GATATGTAATTAATTAATTT-----  447
   |||||||
1067  GATATGTAATTAATTAATTAATTAATTTTTCGCAATTTATGTTAATATCATCT  1126
   |||||||
448  -----AATTCATTTGATGGCAAAATCGTGGCCAAATCTGCTGCCAAAAT  496
   |||||||
1127  CACAACTTCAGAAATTCATTTGATGGCAAAATCGTGGCCAAATCTGCTGCCAAAAT  1186
   |||||||
497  TCGATGAAGGAAGGCTCTCCTAAGCAACATCGTTTCAAAAGTTGACAAACGCTTTGAAC  556
```

```
|||||
1187  TCGATGAAGGAAGGCTCTCCTAAGCAACATCGTTTCAAAAGTTGACAAACGCTTTGAAC  1246
   |||||||
557  TATTAATATTTATCCAAATCTACTGAAATATACCGGTAGCGGCAATTCAGAAATATGATGAC  616
   |||||||
1247  TATTAATATTTATCCAAATCTACTGAAATATACCGGTAGCGGCAATTCAGAAATATGATGAC  1306
   |||||||
617  GACGATGATATGATATCTATCAAAAAGAAAGCTTTCCGAATGCTTCAGATCAGATG  676
   |||||||
1307  GACGATGATGATATGATATCTATCAAAAAGAAAGCTTTCCGAATGCTTCAGATCAGATG  1366
   |||||||
677  ACAGACAGCAAGCTCTTATGCTCTGATCTCAATATTTGATGAGAGAGGAAGCTTTCTTCA  736
   |||||||
1367  ACAGACAGCAAGCTCTTATGCTCTGATCTCAATATTTGATGAGAGAGGAAGCTTTCTTCA  1426
   |||||||
737  TATTTCTGTTGATTTTGCACAGCTCAATGATATCAATTAATTCGAAAAAGAAATATTA  796
   |||||||
1427  TATTTCTGTTGATTTTGCACAGCTCAATGATATCAATTAATTCGAAAAAGAAATATTA  1486
   |||||||
797  TTGAGACACAAATTAATTTGCTCAGCAATTAATTAATTAATTAATTAATTAATTAATTA  856
   |||||||
1487  TTGAGACACAAATTAATTTGCTCAGCAATTAATTAATTAATTAATTAATTAATTAATTA  1546
   |||||||
857  CAGACAGCTCACAATTTGATCAATTAAGCAAAAGATCCATTTGATCCATGATTTGGAGAA  916
   |||||||
917  AAGTTAAGGTTGACAAAAGTTTGAAGCTCATGACCCCTTGGCTCAGCAATTCATTAAC  976
   |||||||
1607  AAGTTAAGGTTGACAAAAGTTTGAAGCTCATGACCCCTTGGCTCAGCAATTCATTAAC  1666
   |||||||
977  TCCAGCTGCTTCATTTTCATTAATTTTCAAAATCGAAGGATATCTTATTTGGGAATG  1036
   |||||||
1667  TCCAGCTGCTTCATTTTCATTAATTTTCAAAATCGAAGGATATCTTATTTGGGAATG  1726
   |||||||
1037  ATGCTCAGATGACCTTGAAGAGATTTCTCATTCATATCAATTAATTCATTTATGTTCC  1096
   |||||||
1727  ATGCTCAGATGACCTTGAAGAGATTTCTCATTCATATCAATTAATTCATTTATGTTCC  1786
   |||||||
1097  CAGTTGTTATTCGGAAGAAATTAATTTTGAACCTTTCGCCAGATGATTCAAAGGAA  1156
   |||||||
1787  CAGTTGTTATTCGGAAGAAATTAATTTTGAACCTTTCGCCAGATGATTCAAAGGAA  1846
   |||||||
1157  CATTCAGATGGAATTAATTTTGAAGAAATTCGCAAGAAATTCGCAAGAAATTCGCAAG  1216
   |||||||
1847  CATTCAGATGGAATTAATTTTGAAGAAATTCGCAAGAAATTCGCAAGAAATTCGCAAG  1906
   |||||||
1217  TCAAGCAATGCTTCCCAAGAGATTTAGACAAATTTAAGGTATTCGATTTCCAAAC  1276
   |||||||
1907  TCAAGCAATGCTTCCCAAGAGATTTAGACAAATTTAAGGTATTCGATTTCCAAAC  1966
   |||||||
1277  G-----  1277
   |||
1967  GGGTTGTTTATATGTTTGAAGATTTTCACTATTAATTAATTAATTAATTAATTAATTT  2026
   |||
1278  -----GGTAGACTACGCTTGAAGCTGCGTGAAGCTGCGTGAAGCTGCGTGAAGCT  1330
   |||||||
2027  GTTTTAAGGTAGACTACGCTTGAAGCTGCGTGAAGCTGCGTGAAGCTGCGTGAAGCT  2086
   |||||||
1331  CAGCTACAGTGAATCAGTTCAATGAGAACTGATTAATGTCATTAATTCAGCGCTGGAGT  1390
   |||||||
2087  CAGCTACAGTGAATCAGTTCAATGAGAACTGATTAATGTCATTAATTCAGCGCTGGAGT  2146
   |||||||
1391  AAGAATTTGATGACTGATGATGTTG-----  1417
   |||||||
2147  AAGAATTTGATGACTGATGATGTTGAGTATGATTAATCAAGCCGAAGCAATTTTTC  2206
   |||||||
1418  -----ACTCCAGTATTTCTACCGGTAGATGATGATGATGATGATGATGATGATGAT  1456
   |||||||
2207  AATGAATAATCTATCAATTTCAATTCATGATATTTCTACCGGTAGATGATGATGATGAT  2266
   |||||||
1457  ACAGTTATGTTCTACAACTCCGAAAGATACAACTATTTGA  1497
   |||||||
```


Db 2267 ACAGTTATGTTCTACACCTCGGAAAAAGTACACATTGGA 2307

RESULT 7
ABL52520
ID ABL52520 standard; DNA: 2307 BP.
XX ABL52520;
XX
XX 17-JUL-2002 (first entry)
XX
XX C. elegans lin-61(n3635) DNA sequence SEQ ID NO:78.
DE
XX
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KM cell proliferation; nematode; cancer; mutant; gene; ds.
XX
XX Caenorhabditis elegans.
OS
XX WO200194545-A2.
XX
XX 13-DEC-2001.
XX
XX 01-JUN-2001: 2001WO-US17909.
XX
XX 02-JUN-2000: 2000US-208802P.
XX
XX (MASI) MASSACHUSETTS INST TECHNOLOGY.
XX
XX Horvitz HR, Davison EM, Lu X;
XX
XX WPI; 2002-401590/43.
XX
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
XX Disclosure; Page 116; 116pp; English.

The present invention describes a substantially pure nematode C. elegans
(Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
(see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
(see ABB78650), where the polypeptides comprise at least 130, 110, 130
contiguous amino acids of the 386, 322 and 498 amino acid sequences given
in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
CC cytostatic activity, and can be used in gene therapy. (I) is useful for
modulating proliferation of a cell, and for identifying the compound that
modulates cell proliferation. (I) can be used for diagnosing an animal
(preferably, human) for the presence of the cell proliferation disease,
or an increased chance of developing the disease, by measuring lin-8,
lin-56 or lin-61 nucleic acid expression in a sample obtained from the
animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
(I) is also useful for identifying the compound that modulates cell
proliferation. (I) and its mutants are useful for understanding cell
proliferative diseases including cancer, as well as in diagnosing and
treating cell proliferative diseases. (I) is also useful for identifying
tumour suppressors in other species such as mammals and may be used to
identify therapeutic compounds. The present sequence represents a
C. elegans lin-61(n3635) DNA sequence from the present invention.

Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;

Query Match 73.7%; Score 1103; DB 24; Length 2307;
Best Local Similarity 81.3%; Pred. No. 2,2e-283;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

QY 1 ATGCTGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAAACTCGATAG 60
DB 467 ATGCTGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAAACTCGATAG 526

QY 61 ACCTACTTGTGGGAATCCTATTATTCATCGCTTCGAGAAAGAAAACTTTCTTCATCCA 120
DB 527 ACCTACTTGTGGGAATCCTATTATTCATCGCTTCGAGAAAGAAAACTTTCTTCATCCA 586

QY 121 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAGAGTT 180
DB 587 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAGAGTT 646
QY 181 ATC-----TTGAAACA 192
DB 647 ATCGTAGTTTCATATTGTTCTGTAATCGTTTAAAAATCAATTTTGTAGTTCGAAACA 706
QY 193 GTGTCATGATATATGACAAAGACCTCGATTCGATCAAGTCAGATGTTTGACAAAT 252
DB 707 GTGTCATGATATATGACAAAGACCTCGATTCGATTCAGTCAGATGTTTGACAAAT 766
QY 253 GAAAAATTGCGGATACAGAGTTCGCTCATGTTATCGAGCTGACAGAAATTTTGG 312
DB 767 GAAAAATTGCGGATACAGAGTTCGCTCATGTTATCGAGCTGACAGAAATTTTGG 826
QY 313 CTCGAATATTTTATCGAGCATATATGTTTGGTGGCAAA----- 350
DB 827 CTCGAATATTTTATCGAGCATATATGTTTGGTGGCAAAAGTATGAGCTCAGCTTT 886
QY 351 ----- 350
DB 887 TCTACTATTTCTAAATAAATGTTCTGTACATAAATTTGAGAAACATCGATTA 946
QY 351 -----CGCCGCAATGAGTAT 366
DB 947 AAACCTCGAAACATTTGTATATAGTAAATTTGAACATTTTCAGCCCGCAATGAGTAT 1006
QY 367 CCCAATATGATATAATTTGTATATGCTCCGCGCTTGCAATCAAGCAAGATACAAAT 426
DB 1007 CCCAATATGATATAATTTGTATATGCTCCGCGCTTGCAATCAAGCAAGATACAAAT 1066
QY 427 GATATGTAATTTATGTAAT----- 447
DB 1067 GATATGTAATTTATGTAATTTATGTAATTTATGTAATTTATGTAATTTATGTAAT 1126
QY 448 -----AATTCATGATGCGGAAATCGCGCCAAATCGTCTGCTCAAAAT 496
DB 1127 CACAACTTCAAAATTTGATGATGCGGAAATCGCGCCAAATCGTCTGCTCAAAAT 1186
QY 497 TCGATGAAGGAGGCTCTCTTAAGCAAGCATGTTTCAAAAGTTGACAAAGCTTTGAC 556
DB 1187 TCGATGAAGGAGGCTCTCTTAAGCAAGCATGTTTCAAAAGTTGACAAAGCTTTGAC 1246
QY 557 TATTAATTTATTCATTTCTACGAAATACGCTACGCGCAATTCAGAAATTTGTGAC 616
DB 1247 TATTAATTTATTCATTTCTACGAAATACGCTACGCGCAATTCAGAAATTTGTGAC 1306
QY 617 GACGATGAATGATCTATACCAAAAGAACTTTCCGAAATGCTTCCAGATGACAGATG 676
DB 1307 GACGATGAATGATCTATACCAAAAGAACTTTCCGAAATGCTTCCAGATGACAGATG 1366
QY 677 ACGACAGACAGTCTTAACTCTGATCTCAATATTTGGATAGACGAGGAAAGCTTTCA 736
DB 1367 ACGACAGACAGTCTTAACTCTGATCTCAATATTTGGATAGACGAGGAAAGCTTTCA 1426
QY 737 TATTTCTGTGATTTGACAGCTGCAATGATATCACTAATTTGGAAAAAGCAATATA 796
DB 1427 TATTTCTGTGATTTGACAGCTGCAATGATATCACTAATTTGGAAAAAGCAATATA 1486
QY 797 TTGACACACAAATAAATTTGCTCAAGCAATAAATAATGAGAAATTCGAATATGACT 856
DB 1487 TTGACACACAAATAAATTTGCTCAAGCAATAAATAATGAGAAATTCGAATATGACT 1546
QY 857 CAGACAGAGTCACATTTGATCAATATAGCAAAAGATCCATTTGATCCATGATTTGAGAA 916
DB 1547 CAGACAGAGTCACATTTGATCAATATAGCAAAAGATCCATTTGATCCATGATTTGAGAA 1606
QY 917 AAGTTAAGTTGGACAAAAGTTTGAAGTATGACCTGATGACCCCTTGGCTGACCAATTAATPACC 976
DB 1607 AAGTTAAGTTGGACAAAAGTTTGAAGTATGACCTGATGACCCCTTGGCTGACCAATTAATPACC 1666


```

FT      /transl_except= (pos:1748..1750, aa:Val)
FT      /transl_except= (pos:1826..1828, aa:Arg)
FT      /transl_except= (pos:1901..1903, aa:Asn)
XX
XX      WO200011168-A2.
XX
XX      02-MAR-2000.
XX
XX      20-AUG-1999; 99WO-US19052.
XX
XX      21-AUG-1998; 98US-0138132.
XX
XX      (UYPR-) UNIV PRINCETON.
XX
XX      Lemischka I, Moore K;
XX
XX      WPI: 2000-237650/20.
XX      P-PSDB; AAY79183.
XX
XX      Hematopoietic stem cell signaling proteins modulating replication and
XX      differentiation for treating immune system disorders and leukemia -
XX
XX      Claim 10; Page 228-229; 256pp; English.
XX
XX      The present sequence is that of a nucleic acid isolated from
XX      mouse primitive stem cells by a method of the invention. It is an
XX      example of claimed isolated nucleic acids (see AYZ94077-294131) that
XX      are specifically expressed in haematopoietic stem cells (HSCs) and
XX      which encode HSC-specific proteins. The HSCs are especially
XX      primitive HSCs (PHSCs) such as umbilical cord cells, bone marrow
XX      cells and foetal liver cells. The encoded proteins (see AAY79176-93)
XX      are growth factors, transcription factors, splicing factors, capping
XX      factors, transport proteins, translation factors or replication
XX      factors that modulate HSC activity, especially differentiation or
XX      replication. The invention provides a claimed method for
XX      identifying PHSC-specific nucleic acids involving: creating a PHSC
XX      cDNA library and a non-PHSC immune cell library; and subtracting
XX      the 2 libraries. Also claimed are methods: for generating a stem
XX      cell/progenitor cell from PHSCs; for identifying the presence of a
XX      PHSC in a sample; for identifying the presence in a sample of a
XX      compound that modulates HSC activity; for using such a compound to
XX      treat an immune system condition, especially leukaemia; for
XX      introducing exogenous nucleic acid into a HSC; and for ex vivo
XX      expansion of HSCs. Also claimed is a PHSC specifically expressing 1
XX      of the claimed nucleic acids, such as the present sequence.
XX
XX      Sequence 3590 BP; 1101 A; 702 C; 797 G; 990 T; 0 other:
XX
XX      Query Match 3.8%; Score 57.4; DB 21; Length 3590;
XX      Best Local Similarity 55.2%; Pred. No. 7,4e-05;
XX      Matches 112; Conservative 0; Mismatches 91; Indels 0; Gaps 0;
XX
XX      1276 CGGGTAGGAGTACCGCTGAGCTGTCGATGTCGAAATATGATTATTTTCCAGCT 1335
XX      1745 CGGTTGAGATGAAATTTAGAGCTGAGATCTCATGTCAGGACCAAGTTAATATGTGAGCC 1804
XX
XX      1336 ACAGTAAATCACTTCATGGAAGAGCTGATTAATGTCATTTCCAGCGCTGGATGAAGAA 1395
XX      1805 ACAGTACCTCGAATTTATTCACCATCTCTGAGAGATACATTTTATGTTGGAGAAAGAG 1864
XX
XX      1396 TTGTGATGAACGTATGATGTGACCTCCCATGATATTTCTACCGATAGATGGTGAAGCG 1455
XX      1865 TATGACCACTGGGTAGACTGTGAGTCCCTGACCTATACCTGTAGGTTGTCAGTTA 1924
XX
XX      1456 CACAGTTATGTTCTACACCTCC 1478
XX      1925 ACTGATATCACTACAGCCTCC 1947
XX
XX      DB
XX
XX      RESULT 10
XX      ABX08841
XX      ID ABX08841 standard; cDNA: 2564 BP.
XX

```

```

AC      ABX08841;
XX
XX      21-JAN-2003 (first entry)
XX
XX      Angiogenesis-associated human polynucleotide sequence #103.
XX
XX      DE
XX      KW Human; angiogenesis-associated transcript; angiogenesis;
XX      KW angiogenesis-associated disease; cancer; cytostatic; gene therapy;
XX      KW gene; ss.
XX
XX      OS Homo sapiens.
XX
XX      PN WO200279492-A2.
XX
XX      PD 10-OCT-2002.
XX
XX      PF 14-FEB-2002; 2002WO-US04915.
XX
XX      PR 14-FEB-2001; 2001US-0784356.
XX      PR 22-FEB-2001; 2001US-0791390.
XX      PR 19-APR-2001; 2001US-285475P.
XX      PR 03-AUG-2001; 2001US-310025P.
XX      PR 13-NOV-2001; 2001US-350666P.
XX      PR 29-NOV-2001; 2001US-334244P.
XX
XX      PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX      PI Murray R, Glynn R, Watson SR, Aziz N;
XX
XX      DR WPI: 2003-040681/03.
XX      DR P-PSDB; ABU03557.
XX
XX      PT Detecting angiogenesis-associated transcript in a cell for diagnosing
XX      PT and treating cancer by contacting a sample with a polynucleotide that
XX      PT exhibits changes in expression level as a function of time in tissue
XX      PT undergoing angiogenesis.
XX
XX      PS Example 2; Page 274-275; 291pp; English.
XX
XX      CC The present invention relates to methods and compositions for
XX      CC detecting an angiogenesis-associated transcript in a cell in
XX      CC a patient. The method involves contacting a biological sample from
XX      CC the patient with a polynucleotide that selectively hybridises to a
XX      CC sequence at least 80% identical to any of the angiogenesis-associated
XX      CC human polynucleotide sequences given in the specification. These
XX      CC angiogenesis-associated polynucleotide sequences comprise genes that
XX      CC exhibit changes in expression levels as a function of time in tissue
XX      CC undergoing angiogenesis. The method and the polynucleotide sequences
XX      CC of the invention are useful for diagnosing and treating angiogenesis
XX      CC and angiogenesis-associated diseases e.g. cancer. The polynucleotide
XX      CC sequences are also useful in the gene therapy of such disorders. The
XX      CC angiogenesis-associated proteins encoded by the polynucleotide
XX      CC sequences are useful as a vaccine for therapeutic and prophylactic
XX      CC immunisation. ABX08739-ABX08853 represent angiogenesis-associated
XX      CC polynucleotide sequences.
XX
XX      SQ Sequence 2564 BP; 832 A; 482 C; 560 G; 690 T; 0 other:
XX
XX      Query Match 3.7%; Score 55.8; DB 25; Length 2564;
XX      Best Local Similarity 54.7%; Pred. No. 0.00017;
XX      Matches 111; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
XX
XX      1276 CGGGTAGGAGTACCGCTGAGCTGTCGATGTCGAAATATGATTATTTTCCAGCT 1335
XX      1153 CGTGTGAGATGAAATTTAGAGCAAGTATCTCATGAGCCAGCTTTAATATGTGAGCC 1212
XX
XX      1336 ACAGTAAATCACTTCATGGAAGAGCTGATTAATGTCATTTCCAGCGCTGGATGAAGAA 1395
XX      1213 ACAGTAACTCGAATTTATTCATCTCTGAGAGATACATTTTATGATGGAGAAAGAG 1272
XX
XX      1396 TTGTGATGAACGTATGATGTGACCTCCCATGATATTTCTACCGATAGATGGTGAAGCG 1455
XX      1273 TATGATCACTGGGTAGACTGTGAGTCCCTGACCTATACCTGTAGGTTGTCAGTTA 1332
XX      DB
XX

```

OY 1456 CACAGTTATGTTCTACACCTCC 1478
| | | | | | | | | |
Db 1333 ACTGATATCACTACACCTCC 1355

RESULT 11
AAS27361

ID AAS27361 standard; cDNA: 2403 BP.

XX AAS27361;

DT 07-NOV-2001 (first entry)

XX cDNA encoding novel signal transduction pathway protein, Seq ID 396.

DE Neuroprotective; cytoskeletal; dermatological; immunosuppressive; tumour;
XX antiinflammatory; anti-HIV; antibacterial; antiinflammatory; cancer;
KW immune system disorder; rheumatoid arthritis; hepatitis C; blood disorder;
KW organ transplant rejection; infection; hepatitis; Gaucher's disease;
KW sickle cell anaemia; hyperproliferative disorder; Parkinson's disease;
KW neurodegenerative disorder; Alzheimer's disease; chromosomal abnormality;
KW chromosomal abnormality; Down syndrome; ischaemia; renal disorder;
KW cardiovascular; respiratory; wound healing; endocrine; Addison's disease;
KW reproductive system; gastrointestinal; liver disorder; AIDS; ss;
KW acquired immune deficiency syndrome.

XX Homo sapiens.

OS Wo200154733-A1.

XX 02-AUG-2001.

PD 17-JAN-2001; 2001WO-US01312.

XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216647.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226681.
PR 22-AUG-2000; 2000US-0226688.
PR 23-AUG-2000; 2000US-0227182.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.

PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.
PR 08-SEP-2000; 2000US-0231414.
PR 08-SEP-2000; 2000US-0232080.
PR 08-SEP-2000; 2000US-0232081.
PR 12-SEP-2000; 2000US-0232968.
PR 14-SEP-2000; 2000US-0232397.
PR 14-SEP-2000; 2000US-0233398.
PR 14-SEP-2000; 2000US-0233399.
PR 14-SEP-2000; 2000US-0233400.
PR 14-SEP-2000; 2000US-0233401.
PR 14-SEP-2000; 2000US-0233063.
PR 14-SEP-2000; 2000US-0233064.
PR 14-SEP-2000; 2000US-0233065.
PR 21-SEP-2000; 2000US-0234223.
PR 21-SEP-2000; 2000US-0234274.
PR 25-SEP-2000; 2000US-0234997.
PR 25-SEP-2000; 2000US-0234998.
PR 26-SEP-2000; 2000US-0235484.
PR 27-SEP-2000; 2000US-0235834.
PR 27-SEP-2000; 2000US-0235836.
PR 29-SEP-2000; 2000US-0236327.
PR 29-SEP-2000; 2000US-0236367.
PR 29-SEP-2000; 2000US-0236368.
PR 29-SEP-2000; 2000US-0236369.
PR 29-SEP-2000; 2000US-0236370.
PR 02-OCT-2000; 2000US-0236802.
PR 02-OCT-2000; 2000US-0237037.
PR 02-OCT-2000; 2000US-0237038.
PR 02-OCT-2000; 2000US-0237039.
PR 02-OCT-2000; 2000US-0237040.
PR 13-OCT-2000; 2000US-0239933.
PR 13-OCT-2000; 2000US-0239937.
PR 20-OCT-2000; 2000US-0240960.
PR 20-OCT-2000; 2000US-0241221.
PR 20-OCT-2000; 2000US-0241785.
PR 20-OCT-2000; 2000US-0241786.
PR 20-OCT-2000; 2000US-0241787.
PR 20-OCT-2000; 2000US-0241808.
PR 20-OCT-2000; 2000US-0241826.
PR 01-NOV-2000; 2000US-0244617.
PR 08-NOV-2000; 2000US-0246474.
PR 08-NOV-2000; 2000US-0246475.
PR 08-NOV-2000; 2000US-0246476.
PR 08-NOV-2000; 2000US-0246477.
PR 08-NOV-2000; 2000US-0246478.
PR 08-NOV-2000; 2000US-0246523.
PR 08-NOV-2000; 2000US-0246524.
PR 08-NOV-2000; 2000US-0246525.
PR 08-NOV-2000; 2000US-0246526.
PR 08-NOV-2000; 2000US-0246527.
PR 08-NOV-2000; 2000US-0246528.
PR 08-NOV-2000; 2000US-0246532.
PR 08-NOV-2000; 2000US-0246609.
PR 08-NOV-2000; 2000US-0246610.
PR 08-NOV-2000; 2000US-0246611.
PR 08-NOV-2000; 2000US-0246613.
PR 17-NOV-2000; 2000US-0249207.
PR 17-NOV-2000; 2000US-0249208.
PR 17-NOV-2000; 2000US-0249209.
PR 17-NOV-2000; 2000US-0249210.
PR 17-NOV-2000; 2000US-0249211.
PR 17-NOV-2000; 2000US-0249212.
PR 17-NOV-2000; 2000US-0249213.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)


```
Db      508 TTGGGAGGCTTCGGGCTCAGAGTCTTGTCTCATTTTGATGGTGGCGAGGGCCTTTG 567
;
QY      1400 ATGAACGTATGATGTGAGCTCCATGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
;
Db      568 ACTACGTGTCGGCGCTTCGACTCCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACTG 627
;
QY      1460 GTTATGTTCTACAACTCC 1478
;
Db      628 GAGACAACCTGCAGCCTCC 646
;

RESULT 2
US-08-852-153-5
; Sequence 5, Application US/08852153
; Patent No. 5914266
; GENERAL INFORMATION:
; APPLICANT: Randazzo, Filippo
; TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,153
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guth, Joseph H.
; REGISTRATION NUMBER: 31,261
; REFERENCE/DOCKET NUMBER: 1224.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3888
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ. ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3255 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-852-153-5

Query Match      3.6%; Score 53.4; DB 2; Length 3255;
Best Local Similarity 54.3%; Pred. No. 1.9e-05;
Matches 108; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY      1280 TAGGACTACGCCCTTGAAGCTGTCACATGTGTCAAAATCATTTTATTTGCCAGCTACAG 1339
;
Db      848 TTGGGAATGAGCTAGAGCTGTGGACAGAGAACCCTCATTTTCAATTTATTTGCCACTA 907
;
QY      1340 TGAATCAGTTTCATGAGAGCTGATAAATGTCATTTTCGAGCGCTGGGATGAGAAATTTG 1399
;
Db      908 TTGGGGAGGTTTCGGGGCTCAGAGGTCCTTGTCACTTTTGATGGGTGGCGAGGGCCTTTG 967
;
QY      1400 ATGAACGTATGATGTGAGCTCCATGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
;
Db      968 ACTACGTGTCGGCGCTTCGACTCCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACTG 1027
;
QY      1460 GTTATGTTCTACAACTCC 1478
;
Db      1028 GAGACAACCTGCAGCCTCC 1046
;

RESULT 3
```

```
US-08-852-153-3
; Sequence 3, Application US/08852153
; Patent No. 5914266
; GENERAL INFORMATION:
; APPLICANT: Randazzo, Filippo
; TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 94608
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/852,153
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Guth, Joseph H.
; REGISTRATION NUMBER: 31,261
; REFERENCE/DOCKET NUMBER: 1224.006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-3888
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ. ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3327 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; US-08-852-153-3

Query Match      3.6%; Score 53.4; DB 2; Length 3327;
Best Local Similarity 54.3%; Pred. No. 1.9e-05;
Matches 108; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY      1280 TAGGACTACGCCCTTGAAGCTGTCACATGTGTCAAAATCAGTTTATTTGCCAGCTACAG 1339
;
Db      920 TTGGGAATGAGCTAGAGCTGTGGACAGAGAACCCTCATTTTCAATTTGCTCCAGCCTA 979
;
QY      1340 TGAATCAGTTTCATGAGAGCTGATAAATGTCATTTTCGAGCGCTGGGATGAGAAATTTG 1399
;
Db      980 TTGGGAGGTTTCGGGGCTCAGAGGTCCTTGTCACTTTTGATGGGTGGCGAGGGCCTTTG 1039
;
QY      1400 ATGAACGTATGATGTGAGCTCCATGATATTTCTACCGATAGATGTGTGAAGCGCACA 1459
;
Db      1040 ACTACGTGTCGGCGCTTCGACTCCCGAGACATCTTCCCTGTGGGCTGTCTTCTTGACTG 1099
;
QY      1460 GTTATGTTCTACAACTCC 1478
;
Db      1100 GAGACAACCTGCAGCCTCC 1118
;

RESULT 4
US-09-620-312D-622
; Sequence 622, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
```

APPLICANT: Xue, Aidong J.
APPLICANT: Yang, Yonghong
APPLICANT: Wang, Jian-Rui
APPLICANT: Zhou, Ping
APPLICANT: Ma, Yungling
APPLICANT: Wang, Dunrui
APPLICANT: Wang, Zhiwei
APPLICANT: John Tillinghast
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: No. 656662el Nucleic Acids and
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 784CIP2B
CURRENT APPLICATION NUMBER: US/09/620,312D
CURRENT FILING DATE: 2000-07-19
PRIOR APPLICATION NUMBER: 09/552,317
PRIOR FILING DATE: 2000-04-25
PRIOR APPLICATION NUMBER: 09/488,725
PRIOR FILING DATE: 2000-01-21
NUMBER OF SEQ ID NOS: 1105
SOFTWARE: PL-Fl-genes Version 1.0
SEQ ID NO 622
LENGTH: 2781
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (263)..(1750)
US-09-620-312D-622

Query Match 3.4%; Score 50.6; DB 4; Length 2781;
Best Local Similarity 53.2%; Pred. No. 0.00011;
Matches 107; Conservative 0; Mismatches 94; Indels 0; Gaps 0;

QY 1278 GGTAGAGCTACGCTTGAAGCTGCTGACATGTGTGAATAATCAATTTATTTGTCCAGTAC 1337
DB 1243 GGTGGGCGATGAAGCTGGAGCCGTGACCTGATGAGACCCCGCTCATCTGTGTGGCCAC 1302
QY 1338 AGTGAATCAGTTCATGAGACGTGATTAATGTCAATTTGACGGCTGGGATGAAGATT 1397
DB 1303 GGTGAAGAGAGTGTGATGAGCTGCTCCACATCCACTTTGACGGCTGGGACAGCGAGTA 1362
QY 1398 TGATGACACTGTATGATGTGACGTCCCATGATTAATTCACGATAGAGATGTGTGAAGCCA 1457
DB 1363 CGACCGAGTGGGTGAGTCCGACAGTCCCGACATCTACCCCGTGGCTGGTGTGAGCTCAC 1422
QY 1458 CAGTTATGTTCTTACAACCTCC 1478
DB 1423 CGGCTACCGAGCTCCAGCTCC 1443

RESULT 5
US-08-852-153-7
Sequence 7, Application US/08852153
Patent No. 5914266
GENERAL INFORMATION:
APPLICANT: Randazzo, Filippo
TITLE OF INVENTION: Mammalian Sex Comb on Midleg Acts as a Tumor Suppressor
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Chiron Corporation
STREET: 4560 Horton Street
CITY: Emeryville
STATE: California
COUNTRY: U.S.A.
ZIP: 94608
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/852,153
FILING DATE:

CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Guth, Joseph H.
REGISTRATION NUMBER: 31,261
REFERENCE/DOCKET NUMBER: 1224, 006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (510) 923-3888
TELEFAX: (510) 655-3542
INFORMATION FOR SRO ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3065 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-852-153-7

Query Match 3.4%; Score 50.2; DB 2; Length 3065;
Best Local Similarity 53.3%; Pred. No. 0.00016;
Matches 106; Conservative 0; Mismatches 93; Indels 0; Gaps 0;

QY 1280 TAGACCTACGCCCTTGAAGCTGCTGACATGTGTGAATAATCAATTTATTTGTCCAGTAC 1339
DB 864 TGGGAATGAAGTTAGAGCTGTAGACAGAAAGACCCCTCATTTTCATTTGCCAGCGACTA 923
QY 1340 TGAATCAGTTCATGAGAGACTGATTAATGTCAATTTGACGGCTGGGATGAAGATTG 1399
DB 924 TTGGAGAGATTCGAGGGCGGAGAGTGTACCTACCTTTGATGGGTGGCGAGCGCATTTG 983
QY 1400 ATGAACCTGATGATGTGACTCCCATGATTAATTCACGATAGATGTGTGAAGCGACA 1459
DB 984 ACTACTGTGCTCCGCTTGTACTCCCGGAGACATCTTCTGTGGCTGGTGTCTTGTGACTG 1043
QY 1460 GTATATGTTCTACAACCTCC 1478
DB 1044 GAGATACCTGCGAGCCACC 1062

RESULT 6
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU

```
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpc-fls
US-08-232-463-14
```

Query Match 2.8%; Score 41.6; DB 1; Length 7218;
Best Local Similarity 2.8%; Pred. No. 0.068;
Matches 11; Conservative 213; Mismatches 162; Indels 0; Gaps 0;

```
QY 580 GAATACGCGTACGCGAATTCAAGAAATGTGGACGACGAATGATGATCTATCACA 639
D 1429 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 640 AAGAAAGACTTTCOCGATGCTCCAGATGACAGATGACAGACAGACTTACGCT 699
D 1369 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 700 GGATCTCAATATGATAGACGAGGAGCTTCTCATATTTCTGTTGATTGACGA 759
D 1309 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 760 GTCAATGATATCACTAAATGCGAAAGAATATATTGACACACAAATAAATGCT 819
D 1249 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 820 CAAGCAATAAATAATGAGAAATCCAGATGATGACTGACGACGACATTTGATCAA 879
D 1189 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 880 TTAGCAAGAATCCATTTGATCCCATGATTGGAGAAAGTTAAGTTGACAAAAGTT 939
D 1129 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR
QY 940 GAGCTCATGACCCCTGGCTGACGA 965
D 1069 RRRATCGAAGCTCCGACGCTGCA 1044
```

RESULT 7

```
US-09-134-001C-212
Sequence 212, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 212
LENGTH: 1284
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-212
```

Query Match 2.8%; Score 41.2; DB 4; Length 1284;
Best Local Similarity 47.3%; Pred. No. 0.04;
Matches 124; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

```
QY 677 ACACAGACAACTCTTTAGCTCTGATCTCAATATTGGATAGACGAGGAAGCTTCTCA 736
D 833 ATGAAAGCAAAATCAGATAGATTGATTTAAATAAATGATTAACGCTATAGAAATGCA 892
QY 737 TATTCCTTGATTTGACGAGTCATAGATGATATCAATCAATGCAAAAAAGAAATTA 796
D 893 ATTGACAGGTGATAGACAAAGTAAAGAAACCAATTAATGTTATGATGCTGCTC 952
QY 797 TTGAGCACAATAAATAATGCTCAAGCAATTAATAAATGAGAAAATCCAGATATGACT 856
D 953 ATATATATGAGATGATGATGCTGATGACACAAATAGACATTTATATGAGAGATTA 1012
QY 857 CAGACAGCTCACTTTGATTCATTAAGCAAAAGATCCATTTATCCATGATTGGAGA 916
D 1013 AATGTATATTTTATTCTTGCATAATTAAGAAAGCAATTCATGATGATTAACAAAT 1072
QY 917 AAGTTAAGTTGACAAAGTT 938
D 1073 TAAATGATATTGATCAAAATTT 1094
```

RESULT 8

```
US-09-134-001C-1557
Sequence 1557, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
PRIOR FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1557
LENGTH: 1890
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1557
```

Query Match 2.7%; Score 40.4; DB 4; Length 1890;
Best Local Similarity 54.8%; Pred. No. 0.082;
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

```
QY 698 CTGATCTCAATATTGATAGACGAGGAAGCTTCTCATATTTCTGTTGATTGACAG 757
D 1589 CGGAATGCGAAGATGCTTGAAGATGGAACCTTATCAAAATTAAGTGTGCTTTTCTA 1648
QY 758 CAGTCAATGATATCACTAAATGCGAAAAGCAATATATTGACACACAAATTAATG 817
D 1649 GAAATCTGATATAAAGATGATGTCACAAATTAATGTAGAAAATAGTGAACAATTTA 1708
QY 818 CTCAGCAATAAATAATGAGAAAT 843
D 1709 ATCGATGATGAAATGCGGCTACT 1734
```

RESULT 9

```
US-08-916-421B-1
Sequence 1, Application US/08916421B
Patent No. 6503729
GENERAL INFORMATION:
APPLICANT: Built et al.
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methano
FILE REFERENCE: PB275
CURRENT APPLICATION NUMBER: US/08/916,421B
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: US 60/024,428
```

```

: PRIOR FILING DATE: 1996-08-22
: NUMBER OF SEQ ID NOS: 3
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 1
: LENGTH: 1664976
: TYPE: DNA
: ORGANISM: Methanococcus jannaschii
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (28222)..(28222)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (28257)..(28258)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84773)..(84773)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84808)..(84808)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (84812)..(84812)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98120)..(98120)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98159)..(98159)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98239)..(98239)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98266)..(98266)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (98343)..(98343)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (10398)..(10398)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (148948)..(148948)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (163385)..(163385)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (191989)..(191989)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (191995)..(191995)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (231980)..(231980)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234187)..(234187)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234220)..(234220)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (234814)..(234814)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (309398)..(309398)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (309418)..(309418)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (312837)..(312837)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (312993)..(312993)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (319226)..(319226)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (559167)..(559167)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (559241)..(559241)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (600992)..(600992)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (622708)..(622708)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (657081)..(657081)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (657203)..(657203)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (674435)..(674435)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (682442)..(682442)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (713652)..(713652)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (741684)..(741684)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (779455)..(779455)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (779676)..(779676)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (855539)..(855539)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (871619)..(871619)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1084830)..(1084830)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1096846)..(1096846)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1119881)..(1119881)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1130881)..(1130881)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1310988)..(1310988)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1313224)..(1313224)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1349473)..(1349473)
: OTHER INFORMATION: n equals a, t, c, or g
: NAME/KEY: misc_feature
: LOCATION: (1349491)..(1349491)
: OTHER INFORMATION: n equals a, t, c, or g
```

```
NAME/KEY: misc-feature
LOCATION: (1470091)..(1470091)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc-feature
LOCATION: (1569020)..(1569020)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc-feature
LOCATION: (1602912)..(1602912)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc-feature
LOCATION: (1603734)..(1603734)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc-feature
LOCATION: (1637998)..(1637998)
OTHER INFORMATION: n equals a, t, c, or g
NAME/KEY: misc-feature
LOCATION: (1664854)..(1664855)
OTHER INFORMATION: n equals a, t, c, or g
US-08-916-421B-1
```

Query Match 2.6%; Score 39; DB 4; Length 1664976;
Best Local Similarity 46.0%; Pred. No. 4.4;
Matches 132; Conservative 0; Mismatches 155; Indels 0; Gaps 0;

```
OY 48 AAAACGATTAAGACCTACTTGTGGGAATCCTATTACATCAGTTCCGAAAGCAAAAC 107
DB 107526 AAAAAGCTTAAAGCTTACTTCTTCATATACATTAATCTGCAAACTCAAGAGATC 107585
OY 108 TTCTTCAATCCAGTGAAGATTCATCATCTTACCTTACAGTATTAATTAAGATTCGAT 167
DB 107586 TACCAATGATGATATGTGACATTTTGAGGTTCCGTTAAAGTTGATATCTAACGAGAT 107645
OY 168 GAAGGAGAGAGATTATCTTGAAGAGTGCATGATTAATGCAAGAACTGGATTCGAT 227
DB 107646 AACATAAGAGATTAAAGTTAAATTTACGCTATGATGATGATGATTTCTAATTTT 107705
OY 228 TCAAGTCAGATGCTTGGACCAATGCAAAAGTTGCGGATACAGATTCGCGCAGT 287
DB 107706 TCTTTTATTTATTAATCACTCAATTAATCAATTAATTAATTAATTAATTTTTCAT 107765
OY 288 TATCGAGCTGACACAGAAATTTGGCTCATATTTTATCGAGATTA 334
DB 107766 CTGAAAAAATTAATTTATGAAAAAATTAATTTTTCATTAATAA 107812
```

RESULT 10
US-09-016-434-237
Sequence 237, Application US/09016434
Patent No. 6500938
GENERAL INFORMATION:
APPLICANT: Janice Au-Young
APPLICANT: Jeffrey J. Seilhamer
TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF SIGNALING
TITLE OF INVENTION: PATHWAY GENE EXPRESSION
NUMBER OF SEQUENCES: 1490
CORRESPONDENCE ADDRESS:
ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/016,434
FILING DATE: HERewith
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:

```
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0002 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 237:
SEQUENCE CHARACTERISTICS:
LENGTH: 216 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: LINDMOTO3
CLONE: 1577179
US-09-016-434-237
```

Query Match 2.6%; Score 38.6; DB 4; Length 216;
Best Local Similarity 58.1%; Pred. No. 0.1;
Matches 68; Conservative 0; Mismatches 49; Indels 0; Gaps 0;

```
OY 1367 ATGTCAATTTGACGCGCTGGGATGAGAAATTTGATGACCTGTATGATGCTGCCATG 1426
DB 5 ATATCACTTTGATGCTGGAGTGAGGCTTTTGATTTACTGTGCGATGATATTTCTCGAG 64
OY 1427 ATATCTACCGATAGATGATGCTGTAAGCGCACAGTATATGTTCTACAACTCCGAAA 1483
DB 65 ATATTTTCCAGCTGGGTGGTGTCCCTGACAGAGATGTATTACAAACCCACAGAA 121
```

RESULT 11
US-09-601-198-138/c
Sequence 138, Application US/09601198
Patent No. 6531583
GENERAL INFORMATION:
APPLICANT: Cassell, Gail H.
APPLICANT: Chen, Ellison Y.
APPLICANT: Glass, Jennifer S.
APPLICANT: Glass, John I.
APPLICANT: Heiner, Cheryl R.
APPLICANT: Lefkowitz, Elliot
TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREA PLASMA
TITLE OF INVENTION: UREA PLASMA
FILE REFERENCE: UAB-13452/22
CURRENT APPLICATION NUMBER: US/09/601,198
CURRENT FILING DATE: 2000-12-08
PRIOR APPLICATION NUMBER: 60/073,189
PRIOR FILING DATE: 1998-01-30
NUMBER OF SEQ ID NOS: 181
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 138
LENGTH: 747
TYPE: DNA
ORGANISM: Ureaplasma urealyticum
US-09-601-198-138

Query Match 2.6%; Score 38.2; DB 4; Length 747;
Best Local Similarity 58.3%; Pred. No. 0.23;
Matches 67; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

```
OY 786 AAGGATATATTTGAGCACACAAATTAATTTCTCAAGCAATTAATAATGGAGAAATCC 845
DB 607 AAAAGAAATTTGCTAATGATTAACAAATATGATGATGATTTTAAAGTAATAATCAA 548
OY 846 AAGATATGACGACGACGACGACATTTGATCAATTAAGCAAAAGATCCATTTGAT 900
DB 547 TAAAGCGGATCAAAATTAACCTACTGATTATATTAATTAAGAAATCCAAAGAT 493
```

RESULT 12

```
US-09-601-198-127/c
; Sequence 127, Application US/09601198
; Patent No. 6531583
; GENERAL INFORMATION:
; APPLICANT: Cassell, Gall H.
; APPLICANT: Chen, Ellison Y.
; APPLICANT: Glass, Jennifer S.
; APPLICANT: Glass, John I.
; APPLICANT: Helner, Cheryl R.
; APPLICANT: Lefkowitz, Elliot
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND METHOD FOR DETECTING UREAPLASMA
; FILE REFERENCE: UAB-13452/22
; CURRENT APPLICATION NUMBER: US/09/601,198
; CURRENT FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/073,189
; PRIOR FILING DATE: 1998-01-30
; NUMBER OF SEQ ID NOS: 181
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 127
; LENGTH: 972
; TYPE: DNA
; ORGANISM: Ureaplasma urealyticum
US-09-601-198-127

Query Match          2.5%; Score 37.2; DB 4; Length 972;
Best Local Similarity 54.3%; Pred. No. 0.5;
Matches 75; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 790 GAATATATTGACGACACCAATTAATTCGTCAGACATATAAAGAGAGAAATCCAGCA 849
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 798 GAATTAATTAAGATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 739
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 850 TATGACTCGAGAGACATCAATTTGATTCATTAAGCAAAAGATCCATTCATCCAGATT 909
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 738 TATGATTTATGATGTCACGATGATGATGATGATGATGATGATGATGATGATGAT 679
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 910 TGGAGAAAAGTTAAGTT 927
    | ||||| |
DB 678 CGTTATCATTTTAAAGGT 661
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 13
US-09-205-258-243
; Sequence 243, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007P1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
```

```
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 243
; LENGTH: 2271
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (553)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2267)
; OTHER INFORMATION: n equals a,t,g, or c
```



```
|||||
Db 181 ATCTCGAAACAGGTCATGATTAATGACAAAGCTCGATTCGATTCAGTACAGATGG 240
Oy 241 TTTCAGCAATTTGAAAAAGTTGGGATGACAGAGTTCGGCTCAAGTTATGGAGCTGAC 300
Db 241 TTTCAGCAATTTGAAAAAGTTGGGATGACAGAGTTCGGCTCAAGTTATGGAGCTGAC 300
Oy 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTTGGTTGGCAACGCCGCAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTTGGTTGGCAACGCCGCAATG 360
Oy 361 AGTATCCCATATGATGATAAAATGATATGCTCCGCCGCTTGCAATCAAGAGAAATAC 420
Db 361 AGTATCCCATATGATGATAAAATGATATGCTCCGCCGCTTGCAATCAAGAGAAATAC 420
Oy 421 CAAATGATATGATAAATTTATTAATTTGCAATTTGATGGCGAAATCGTCGCCCAACT 480
Db 421 CAAATGATATGATAAATTTATGTAATTAATTTGCAATTTGATGGCGAAATCGTCGCCCAACT 480
Oy 481 TCGCTGCTCCAAAATTTGATGAAGGAGGCTCTCCTTAAGCAACATCGTTTCAAAAGT 540
Db 481 TCGCTGCTCCAAAATTTGATGAAGGAGGCTCTCCTTAAGCAACATCGTTTCAAAAGT 540
Oy 541 GGACAAGCTCTTGAACCTTAATTTATTTCAATTTCTACTGAAATAGCGTACGCGAAT 600
Db 541 GGACAAGCTCTTGAACCTTAATTTATTTCAATTTCTACTGAAATAGCGTACGCGAAT 600
Oy 601 CAAGAAATATGTCGACGACGATGATATCTATCACAAGAAAGATTTCCGCAATCG 660
Db 601 CAAGAAATATGTCGACGACGATGATATCTATCACAAGAAAGATTTCCGCAATCG 660
Oy 661 CTTCAGATGACATGACGACGACAGACAGCTTTAGCTCTGATTCATTAATTTGATAGC 720
Db 661 CTTCAGATGACATGACGACGACAGACAGCTTTAGCTCTGATTCATTAATTTGATAGC 720
Oy 721 GAGGAAAGCTCTTCAATTTCTGATTTGAGATTTGAGAGTCAATGATCAACTAAT 780
Db 721 GAGGAAAGCTCTTCAATTTCTGATTTGAGATTTGAGAGTCAATGATCAACTAAT 780
Oy 781 GCGAAAAAGAAATATTTGAGACACACAATAAATTTGCTCAAGCAATTTAAAAATGAGAA 840
Db 781 GCGAAAAAGAAATATTTGAGACACACAATAAATTTGCTCAAGCAATTTAAAAATGAGAA 840
Oy 841 AATCCAGATATGACTGACGACGCTCATTTGATCAATTAGCAAAAGATTCATTTGAT 900
Db 841 AATCCAGATATGACTGACGACGCTCATTTGATCAATTAGCAAAAGATTCATTTGAT 900
Oy 901 CCCATGATTTGGAGAAAGTTAAGTTGAGCAAAAAGTTGAGCTCATGCCCTTGCT 960
Db 901 CCCATGATTTGGAGAAAGTTAAGTTGAGCAAAAAGTTGAGCTCATGCCCTTGCT 960
Oy 961 CAGCAATTCATTAACCTTCACGCTGCTTCGATTCCTCAAAATTTGCAAAAGATTCAGAT 1020
Db 961 CAGCAATTCATTAACCTTCACGCTGCTTCGATTCCTCAAAATTTGCAAAAGATTCAGAT 1020
Oy 1021 CTATATGTCGAAATGATGATGTCACAGATGCTTGAAGACAGTTTCTATTCATTCAT 1080
Db 1021 CTATATGTCGAAATGATGATGTCACAGATGCTTGAAGACAGTTTCTATTCATTCAT 1080
Oy 1081 AATCATTTATGTTCCAGTTGGTTATGCGGAAAAAGTTAATTTGGAAGCTTGTCGCCA 1140
Db 1081 AATCATTTATGTTCCAGTTGGTTATGCGGAAAAAGTTAATTTGGAAGCTTGTCGCCA 1140
Oy 1141 GATGAGTTCAAGAGCAATTCAGATGAGATTAATTTGAGAAAGATTCGCAAGAAC 1200
Db 1141 GATGAGTTCAAGAGCAATTCAGATGAGATTAATTTGAGAAAGATTCGCAAGAAC 1200
Oy 1201 CTACCGCTTGAATGTTCAAGCTTCCTCCAAAGAGATTTAGCAAAATTTAAGSTA 1260
Db 1201 CTACCGCTTGAATGTTCAAGCTTCCTCCAAAGAGATTTAGCAAAATTTAAGSTA 1260
Oy 1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAGAGTGTGACATGTGGAATAATAG 1320
Db 1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAGAGTGTGACATGTGGAATAATAG 1320
```

```
Db 1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAGAGTGTGACATGTGGAATAATAG 1320
Oy 1321 TTATTTGTCAGCTACAGTAAATTCAGTTTCATGGAAGACGATGAATATGCAATTTGCAC 1380
Db 1321 TTATTTGTCAGCTACAGTAAATTCAGTTTCATGGAAGACGATGAATATGCAATTTGCAC 1380
Oy 1381 GCGTGGATGAAAGATTTGATGAACCTGATGATGTCGACCTCCATGATATTCACCGATA 1440
Db 1381 GCGTGGATGAAAGATTTGATGAACCTGATGATGTCGACCTCCATGATATTCACCGATA 1440
Oy 1441 GGATGCTGTGAAGGCGACAGTTATGTTCTACACCTCCGAAAAAGTCAACTATTGA 1497
Db 1441 GGATGCTGTGAAGGCGACAGTTATGTTCTACACCTCCGAAAAAGTCAACTATTGA 1497
```

RESULT 2

```
US-09-872-523-73
; Sequence 73, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-73
```

```
Query Match 99.9%; Score 1495.4; DB 10; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Oy 1 ATGTCGAAATTTGAAAAATTTGTCAGAGCTTAACAAAAATGCGACAGAAAATCGATTAAG 60
Db 1 ATGTCGAAATTTGAAAAATTTGTCAGAGCTTAACAAAAATGCGACAGAAAATCGATTAAG 60
Oy 61 ACCTACTTGTGGGAATCCTATTTCATCAGTTCGAGAAAGGAAAAAATCTTTTCATTCCA 120
Db 61 ACCTACTTGTGGGAATCCTATTTCATCAGTTCGAGAAAGGAAAAAATCTTTTCATTCCA 120
Oy 121 GTTGAGCATTCATCGTAACCTTAACAGTTAATTTTAAAGAAATGCGTGAAGAGAGTT 180
Db 121 GTTGAGCATTCATCGTAACCTTAACAGTTAATTTTAAAGAAATGCGTGAAGAGAGTT 180
Oy 181 ATCTTCGAAACAGGTCATGATTAATGACAAGAACTGCGATTTGATTCATGATGAGT 240
Db 181 ATCTTCGAAACAGGTCATGATTAATGACAAGAACTGCGATTTGATTCATGATGAGT 240
Oy 241 TTTCAGCAATTTGAAAAAGTTTGGGATGACAGAGTTCGCTCAAGTTATGCGAGCTGAC 300
Db 241 TTTCAGCAATTTGAAAAAGTTTGGGATGACAGAGTTCGCTCAAGTTATGCGAGCTGAC 300
Oy 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTTGGTTGGCAACGCCGCAATG 360
Db 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTTGGTTGGCAACGCCGCAATG 360
Oy 361 AGTATCCCATATGATGATAAATGATATGCTCCGCCGCTTGCAATCAAGAGAAATAC 420
Db 361 AGTATCCCATATGATGATAAATGATATGCTCCGCCGCTTGCAATCAAGAGAAATAC 420
Oy 421 CAAATGATATGATAAATTTATGTAATTAATTTGCAATTTGATGGCGAAATCGTCGCCCAACT 480
Db 421 CAAATGATATGATAAATTTATGTAATTAATTTGCAATTTGATGGCGAAATCGTCGCCCAACT 480
```

```

OY 481 TCGCTGTCTCCAAAATTCGATGAAGGAGGCTCTCTCAAGCAGCATGTTCAAGTT 540
    |||||||
DB 481 TCGCTGTCTCCAAAATTCGATGAAGGAGGCTCTCTCAAGCAGCATGTTCAAGTT 540
OY 541 GGACACAGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
    |||||||
DB 541 GGACACAGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
OY 601 CAAGAAATATGTGAGCAGCAGATGAATGATATCAACAAAGAAAGACTTCCCAATG 660
    |||||||
DB 601 CAAGAAATATGTGAGCAGCAGATGAATGATATCAACAAAGAAAGACTTCCCAATG 660
OY 661 CTTCAGATGACAGATGACAGACAGCAAGCTTTAGCTCTGATCTCAATATTTGATAG 720
    |||||||
DB 661 CTTCAGATGACAGATGACAGACAGCAAGCTTTAGCTCTGATCTCAATATTTGATAG 720
OY 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGCAGCAGTCATGATATCACTAAT 780
    |||||||
DB 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGCAGCAGTCATGATATCACTAAT 780
OY 781 GCGAAAAAGGATATATGAGCAGCAACAATAAATTTGCTCAAGCAATAAAGATGAGAA 840
    |||||||
DB 781 GCGAAAAAGGATATATGAGCAGCAACAATAAATTTGCTCAAGCAATAAAGATGAGAA 840
OY 841 AATCCAGATATGACTCAGACAGCAGTCACATTTGATCAATTAAGCAAAAGATCCAAATGAT 900
    |||||||
DB 841 AATCCAGATATGACTCAGACAGCAGTCACATTTGATCAATTAAGCAAAAGATCCAAATGAT 900
OY 901 CCCATGATTTGAGAAAAAGTTAAGTTGACAAAAAGTTGAGCTCATGCCCTTGCGCT 960
    |||||||
DB 901 CCCATGATTTGAGAAAAAGTTAAGTTGACAAAAAGTTGAGCTCATGCCCTTGCGCT 960
OY 961 CAGCAATTCATTAACCTCCAGCTGCTTCGATTTCCAAATTTGCAAAACGTAAGATAT 1020
    |||||||
DB 961 CAGCAATTCATTAACCTCCAGCTGCTTCGATTTCCAAATTTGCAAAACGTAAGATAT 1020
OY 1021 CTATATGTGGAATGATGATGCTCAGATGCAATGCAAGCACTTTCTATATATCAAT 1080
    |||||||
DB 1021 CTATATGTGGAATGATGATGCTCAGATGCAATGCAAGCACTTTCTATATATCAAT 1080
OY 1081 AATACATTTATTTTCCAGTTGTTATGCGAAAAAGTAAATTTGGAAGCTTTGCGCCA 1140
    |||||||
DB 1081 AATACATTTATTTTCCAGTTGTTATGCGAAAAAGTAAATTTGGAAGCTTTGCGCCA 1140
OY 1141 GATGAGTTCAAGAAACATTCAGATGAGGATGGAATCTTGAGAGAAAGATCCGAGAAC 1200
    |||||||
DB 1141 GATGAGTTCAAGAAACATTCAGATGAGGATGGAATCTTGAGAGAAAGATCCGAGAAC 1200
OY 1201 CTACCGCTTGACTTGTTCAGGCCAATGCCCTTCCCAAGAGATTAAGCAAAATTTAAGTA 1260
    |||||||
DB 1201 CTACCGCTTGACTTGTTCAGGCCAATGCCCTTCCCAAGAGATTAAGCAAAATTTAAGTA 1260
OY 1261 AATTCGATTTCCAAAACGGGTAGAGCTACGCCCTTGAAGCTGTCGATGTCATAAATCAG 1320
    |||||||
DB 1261 AATTCGATTTCCAAAACGGGTAGAGCTACGCCCTTGAAGCTGTCGATGTCATAAATCAG 1320
OY 1321 TTTATTTTGCAGCTACAGTAAATCAAGTTCATGGAAGCTGATTAATTCATTTGCGAC 1380
    |||||||
DB 1321 TTTATTTTGCAGCTACAGTAAATCAAGTTCATGGAAGCTGATTAATTCATTTGCGAC 1380
OY 1381 GGCTGGAGTGAAGAAATTTGATGAAGTATGATGAGTATGAGTATTCACCGATA 1440
    |||||||
DB 1381 GGCTGGAGTGAAGAAATTTGATGAAGTATGATGAGTATGAGTATTCACCGATA 1440
OY 1441 GGATGCTGTGAAGCGCAGAGTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 1497
    |||||||
DB 1441 GGATGCTGTGAAGCGCAGAGTATGTTCTACAACTCCGAAAAAGTACAACTATTGA 1497

```

RESULT 3

US-09-872-523-74

; Sequence 74, Application US/09872523
; Patent No. US20020137906A1

```

; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 74
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
; US-09-872-523-74

Query Match          99.9%; Score 1495.4; DB 10; Length 1497;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTAAACAAAAATTCGACAGAAAACCTCATTAAG 60
    |||||||
DB 1 ATGCTGAATTTCTGAAAATTTGTCAGAGCTAAACAAAAATTCGACAGAAAACCTCATTAAG 60
OY 61 ACCTACTGTGGGAATCCATTTACATCACTTCGAGAAAGAAAACTTTCATATCA 120
    |||||||
DB 61 ACCTACTGTGGGAATCCATTTACATCACTTCGAGAAAGAAAACTTTCATATCA 120
OY 121 GTTGAAGCATTCATACCTTAACCTTAACTTTTAAAGCAATGCGTGAAGAGAGTT 180
    |||||||
DB 121 GTTGAAGCATTCATACCTTAACCTTAACTTTTAAAGCAATGCGTGAAGAGAGTT 180
OY 181 ATCTTCGAACAGTGTCTCATGATTAATGCAAAAGACTGCGATTCATTAAGTCAGATG 240
    |||||||
DB 181 ATCTTCGAACAGTGTCTCATGATTAATGCAAAAGACTGCGATTCATTAAGTCAGATG 240
OY 241 TTTGACGAAATGAAAAATTTGCGGATACAGATCTGCGTCAGTTATTCGAGCTGAC 300
    |||||||
DB 241 TTTGACGAAATGAAAAATTTGCGGATACAGATCTGCGTCAGTTATTCGAGCTGAC 300
OY 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGTTGGCAAAACCGCAATG 360
    |||||||
DB 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGTTGGCAAAACCGCAATG 360
OY 361 AGTGATCCCAATATGATTAATTTGTAATGCTCGCGCTTGCAATCAACGAATATC 420
    |||||||
DB 361 AGTGATCCCAATATGATTAATTTGTAATGCTCGCGCTTGCAATCAACGAATATC 420
OY 421 CAAATATGATGTAATATGTAATATTTGATGATGAGGAAATGCTGCGCCAAATC 480
    |||||||
DB 421 CAAATATGATGTAATATGTAATATTTGATGATGAGGAAATGCTGCGCCAAATC 480
OY 481 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTCAAGCAAGCATGTTCAAGTT 540
    |||||||
DB 481 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTCAAGCAAGCATGTTCAAGTT 540
OY 541 GGACACAGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
    |||||||
DB 541 GGACACAGCTTGAACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 600
OY 601 CAAGAAATATGTGAGCAGCAGATGAATGATATCAACAAAGAAAGACTTCCCAATG 660
    |||||||
DB 601 CAAGAAATATGTGAGCAGCAGATGAATGATATCAACAAAGAAAGACTTCCCAATG 660
OY 661 CTTCAGATGACAGATGACAGACAGCAAGCTTTAGCTCTGATCTCAATATTTGATAG 720
    |||||||
DB 661 CTTCAGATGACAGATGACAGACAGCAAGCTTTAGCTCTGATCTCAATATTTGATAG 720
OY 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGCAGCAGTCATGATATCACTAAT 780
    |||||||
DB 721 GAGGGAAGCTTCTTCATATTTCTGTTGATTTGCAGCAGTCATGATATCACTAAT 780

```

```
OY      781  GCGAAAAAGAAATATATTTGAGCACACAAATPAAAAATTCCTCAAGCAATPAAAAATGAGAA  840
      |||||||
Db      781  GCGAAAAAGAAATATATTTGAGCACACAAATPAAAAATTCCTCAAGCAATPAAAAATGAGAA  840
OY      841  AATCCAGATATGACTCAGACGACGTCACTTTGATCAATTAGCAAAAAGATCCCAATTGAT  900
      |||||||
Db      841  AATCCAGATATGACTCAGACGACGTCACTTTGATCAATTAGCAAAAAGATCCCAATTGAT  900
OY      901  CCCATGATTTGGAGAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT  960
      |||||||
Db      901  CCCATGATTTGGAGAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT  960
OY      961  CAGCAATTCATAATACCTCCACGTCGCTTGATTCATAATTTTGCAAAACCTGAAGATAT  1020
      |||||||
Db      961  CAGCAATTCATAATACCTCCACGTCGCTTGATTCATAATTTTGCAAAACCTGAAGATAT  1020
OY      1021  CTTATTTGGGAATGGAATGGGCGAGTCACTTGGAACAGTTTCTTATTCATATCAAT  1080
      |||||||
Db      1021  CTTATTTGGGAATGGAATGGGCGAGTCACTTGGAACAGTTTCTTATTCATATCAAT  1080
OY      1081  AATACATTATATGTTCCAGTTGTTATGCGAAAAAGTATATTTGAACTTTGTTCCGCA  1140
      |||||||
Db      1081  AATACATTATATGTTCCAGTTGTTATGCGAAAAAGTATATTTGAACTTTGTTCCGCA  1140
OY      1141  GATGAGTTTCAAAGACATTCAGATGGATGAAATACTTGGAGAAAAGATCTGCAGAAAC  1200
      |||||||
Db      1141  GATGAGTTTCAAAGACATTCAGATGGATGAAATACTTGGAGAAAAGATCTGCAGAAAC  1200
OY      1201  CTACCGTTGACCTGTTCACAGCCAAATGCCCTTCCCAAGAGATAGACAAATTTAAAGTA  1260
      |||||||
Db      1201  CTACCGTTGACCTGTTCACAGCCAAATGCCCTTCCCAAGAGATAGACAAATTTAAAGTA  1260
OY      1261  ATTCTGATTTCCAAACGGGTAGACACTAGCCTTGAACTGCTGACATGTGGAATAACAG  1320
      |||||||
Db      1261  ATTCTGATTTCCAAACGGGTAGACACTAGCCTTGAACTGCTGACATGTGGAATAACAG  1320
OY      1321  TTTATTTTGTCCAGCTACAGTGAATAATCATGTTCAAGGAAGCTGATPAAATGTCATTTGAC  1380
      |||||||
Db      1321  TTTATTTTGTCCAGCTACAGTGAATAATCATGTTCAAGGAAGCTGATPAAATGTCATTTGAC  1380
OY      1381  GCGCGGATGAGAAATTTGATGAACCTGATGATGTGGACCTCCATGATTTCTCCGATA  1440
      |||||||
Db      1381  GCGCGGATGAGAAATTTGATGAACCTGATGATGTGGACCTCCATGATTTCTCCGATA  1440
OY      1441  GGATGCTGTGAAGCGCACAGTTATGTTCTACAACTCCGAAAAAGTACAATATTGA  1497
      |||||||
Db      1441  GGATGCTGTGAAGCGCACAGTTATGTTCTACAACTCCGAAAAAGTACAATATTGA  1497

RESULT 4
US-09-872-523-75
: Sequence 75, Application US/09872523
: Patent No. US20020137906A1
: GENERAL INFORMATION:
: APPLICANT: Horvitz, H. Robert
: APPLICANT: Davison, Ewa M.
: APPLICANT: Lu, Xiaowei
: TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
: FILE REFERENCE: 01997/536002
: CURRENT APPLICATION NUMBER: US/09/872,523
: PRIOR FILING DATE: 2001-06-01
: PRIOR APPLICATION NUMBER: US 60/208,802
: NUMBER OF SEQ ID NOS: 78
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 75
: LENGTH: 1497
: TYPE: DNA
: ORGANISM: Caenorhabditis elegans
US-09-872-523-75
Query Match      99.9%; Score 1495.4; DB 10; Length 1497;
```

```
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1496; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY      1  ATGCTGAATTTTTCGAAAAATTTGTCAGAGCTPACAAAAAATTCGACAGAAAACTCGATPAG  60
      |||||||
Db      1  ATGCTGAATTTTTCGAAAAATTTGTCAGAGCTPACAAAAAATTCGACAGAAAACTCGATPAG  60
OY      61  ACCTACTTGTGGGAATCTATTTACATCACTTCGAGAGAAAAACTTTCTTCATTTCA  120
      |||||||
Db      61  ACCTACTTGTGGGAATCTATTTACATCACTTCGAGAGAAAAACTTTCTTCATTTCA  120
OY      121  GTTGAACCATTTCAATGTAACCTTACAGTTAATTTTAAGCAATGCGTAAGGAAGGAT  180
      |||||||
Db      121  GTTGAACCATTTCAATGTAACCTTACAGTTAATTTTAAGCAATGCGTAAGGAAGGAT  180
OY      181  ATCTTCGAACAGTGTGTCATGATTTATGACAGAACTGCAATTCGATTCAGTCAGATG  240
      |||||||
Db      181  ATCTTCGAACAGTGTGTCATGATTTATGACAGAACTGCAATTCGATTCAGTCAGATG  240
OY      241  TTTGCACGAATTTGAAAAAGTTTGCGGATPACAGAGTTTGCTCAGTTTATTCGAGCTGAC  300
      |||||||
Db      241  TTTGCACGAATTTGAAAAAGTTTGCGGATPACAGAGTTTGCTCAGTTTATTCGAGCTGAC  300
OY      301  ACGAAATTTTGGCTCAATATTTTATGCGAGATATGTTGGTGGCAACGGCCGCAATG  360
      |||||||
Db      301  ACGAAATTTTGGCTCAATATTTTATGCGAGATATGTTGGTGGCAACGGCCGCAATG  360
OY      361  AGTATCCCAATATGATTAATAATTTGATATGCTCCGCGCTTGCAATCAAGAAATPAC  420
      |||||||
Db      361  AGTATCCCAATATGATTAATAATTTGATATGCTCCGCGCTTGCAATCAAGAAATPAC  420
OY      421  CAAATATATATGTAATTTATGTAATPAAATTCATGATGCGCAATTCGCGCAACT  480
      |||||||
Db      421  CAAATATATATGTAATTTATGTAATPAAATTCATGATGCGCAATTCGCGCAACT  480
OY      481  TCGCTGTCTCCAAAATTCGATGGAAGGAGAGCTCTCTAAGCAAGATCGTTTCAAGTT  540
      |||||||
Db      481  TCGCTGTCTCCAAAATTCGATGGAAGGAGAGCTCTCTAAGCAAGATCGTTTCAAGTT  540
OY      541  GGACAAAGCTTGTGAATTAATTAATTTCCAAATTCCTACTGAATPACGCGTACGCAAT  600
      |||||||
Db      541  GGACAAAGCTTGTGAATTAATTAATTTCCAAATTCCTACTGAATPACGCGTACGCAAT  600
OY      601  CAAAGAAATATGTCAGACACGAAATGATATCTATTCACAAAGAAAGCTTTCCGGAATG  660
      |||||||
Db      601  CAAAGAAATATGTCAGACACGAAATGATATCTATTCACAAAGAAAGCTTTCCGGAATG  660
OY      661  CTTCCAGATGAGATGACGACAGCAAGTCTTTAGCTCTGATCTCAATATTTGGATGAC  720
      |||||||
Db      661  CTTCCAGATGAGATGACGACAGCAAGTCTTTAGCTCTGATCTCAATATTTGGATGAC  720
OY      721  GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACAGTTCAGCAATGATATCAACTAAT  780
      |||||||
Db      721  GAGGGAAGCTTCTTCATATTTCTGTTGATTTGACAGTTCAGCAATGATATCAACTAAT  780
OY      781  GCGAAAAAGGAATATATTTGAGCACACAAATPAAAAATTCCTCAAGCAATPAAAAATGAGAA  840
      |||||||
Db      781  GCGAAAAAGGAATATATTTGAGCACACAAATPAAAAATTCCTCAAGCAATPAAAAATGAGAA  840
OY      841  AATCCAGATATGACTCAGACGACGTCACTTTGATCAATTAGCAAAAAGATCCCAATTGAT  900
      |||||||
Db      841  AATCCAGATATGACTCAGACGACGTCACTTTGATCAATTAGCAAAAAGATCCCAATTGAT  900
OY      901  CCCATGATTTGGAGAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT  960
      |||||||
Db      901  CCCATGATTTGGAGAAAAGTTAAGTTGAGCAAAAAGTTTGAGCTCATCGACCCCTTGCT  960
OY      961  CAGCAATTCATAATACCTCCACGTCGCTTGATTCATAATTTTGCAAAACCTGAAGATAT  1020
      |||||||
Db      961  CAGCAATTCATAATACCTCCACGTCGCTTGATTCATAATTTTGCAAAACCTGAAGATAT  1020
OY      1021  CTTATTTGGGAATGGAATGGGCGAGTCACTTGGAACAGTTTCTTATTCATATCAAT  1080
      |||||||
```

Db 1021 CTTATTTGTGGGAATGGATGGATGTCAGATGCACCTTGAAGACAGTTTCCATTCATATCAAT 1080
QY 1081 AATACATTTATGTGTCOCAGTTGGTTATGGGAAAGTAAATTTGGAATTTGTGTCGCCA 1140
Db 1081 AATACATTTATGTGTCOCAGTTGGTTATGGGAAAGTAAATTTGGAATTTGTGTCGCCA 1140
QY 1141 GATGAGTTCAAGGAACATTCAGATGAGATGATACTTTGGAGAAAGATCTCAGAAACC 1200
Db 1141 GATGAGTTCAAGGAACATTCAGATGAGATGATACTTTGGAGAAAGATCTCAGAAACC 1200
QY 1201 CTACCGCTTGACTGTTCAAGCCATGCTTCCCAAGAGATTAGACAAATTTAAGTA 1260
Db 1201 CTACCGCTTGACTGTTCAAGCCATGCTTCCCAAGAGATTAGACAAATTTAAGTA 1260
QY 1261 ATTCTGATTTCCAAACGGGTAGACACTGACCTTGAAAGCTGACATGTGTGAATATCAG 1320
Db 1261 ATTCTGATTTCCAAACGGGTAGACACTGACCTTGAAAGCTGACATGTGTGAATATCAG 1320
QY 1321 TTTATTTTGCACGCTACAGTGAATCAGTTCAATGAGAGACTGATAATTCATATTTGCAC 1380
Db 1321 TTTATTTTGCACGCTACAGTGAATCAGTTCAATGAGAGACTGATAATTCATATTTGCAC 1380
QY 1381 GGCTGGAGTGAAGAAATTTATGATGATGTGATGTGACTCCCATGATATTTCTACCGATA 1440
Db 1381 GGCTGGAGTGAAGAAATTTATGAACTGATGATGTGACTCCCATGATATTTCTACCGATA 1440
QY 1441 GGATGGTGTGAAGCGACAGTTATGTTTCACAACTCCGAAAAGTACAACTATTGA 1497
Db 1441 GGATGGTGTGAAGCGACAGTTATGTTTCACAACTCCGAAAAGTACAACTATTGA 1497

RESULT 5
US-09-872-523-76 : Sequence 76, Application US/09872523
Patent No. US20020137906A1
GENERAL INFORMATION:
APPLICANT: Horvitz, H. Robert
APPLICANT: Davison, Ewa M.
APPLICANT: Lu, Xiaowei
TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
FILE REFERENCE: 01997/536002
CURRENT APPLICATION NUMBER: US/09/872,523
CURRENT FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: US 60/208,802
PRIOR FILING DATE: 2000-06-02
NUMBER OF SEQ ID NOS: 78
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 76
LENGTH: 2307
TYPE: DNA
ORGANISM: Caenorhabditis elegans
US-09-872-523-76

Query Match 73.7%; Score 1103; DB 10; Length 2307;
Best local Similarity 81.3%; Pred. No. 1.3e-282;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

QY 1 ANGTCGATTTCTGAAATTTGTCAGAGCTAACAAAATTCGACAGAAAACCTCCGTAAG 60
Db 467 ANGTCGATTTCTGAAATTTGTCAGAGCTAACAAAATTCGACAGAAAACCTCCGTAAG 526
QY 61 ACCCTACTTGGGAAATCCATTTTACATCAGTTGCGAAGAAAACCTCTTTCATTTCA 120
Db 527 ACCCTACTTGGGAAATCCATTTTACATCAGTTGCGAAGAAAACCTCTTTCATTTCA 586
QY 121 GTTGAAGCAATTCATGTAACCTTACAGATTAATTTTAAAGAAATGCGTGAAGAGAGTT 180
Db 587 GTTGAAGCAATTCATGTAACCTTACAGATTAATTTTAAAGAAATGCGTGAAGAGAGTT 646
QY 181 ATC-----TTGCAATA 192
Db 647 ATCTGAGATTCAATTTGTTGTAATTCGGTTTAAATCAATTTTGTAGTTGCAACA 706

QY 193 GTGTCCATGATTTATGACAAAGACTGCGATTCGATTTCAAGTGCAGATGTTTGCAGAAAT 252
Db 707 GTGTCCATGATTTATGACAAAGAACTGCGATTCGATTTCAAGTGCAGATGTTTGCAGAAAT 766
QY 253 GAAAAGTTTGGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGACAGAAATTTTGG 312
Db 767 GAAAAGTTTGGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGACAGAAATTTTGG 826
QY 313 CTCATATTTTATCGACATATGTTTGGTTGGCAA----- 350
Db 827 CTCATATTTTATCGACATATGTTTGGTTGGCAAAGTAAGTTGAGCTCAGCTCTT 886
QY 351 ----- 350
Db 887 TCTACTATTCTAATAATAATAGTTCTGTACATAAATTCAGAGAAACATTCATTA 946
QY 351 -----CGCCCAATGAGTAT 366
Db 947 AAACCTCGAAACATTTGTAAATAGTAAATTTGAACATTTTCAGCCGCCCAATGAGTGAT 1006
QY 367 CCCAATATGATPAAATTTGATATGCTCCGCCGCTTGCAATGACGAAAGATPACCAAAAT 426
Db 1007 CCCAATATGATPAAATTTGATATGCTCCGCCGCTTGCAATGACGAAAGATPACCAAAAT 1066
QY 427 GATATGTAATTTATGTAAT----- 447
Db 1067 GATATGTAATTTATGTAATGTAATGTTGTTTTTCCGAATTTATGTTAATATCATCT 1126
QY 448 -----AATTCATTTGATGGCAATTCGTGGCCAAACTTGCCTGCTCCAAAT 496
Db 1127 CACAACTCGAAATTCATTTGATGATGCGAAATGCTGCGCCAAACTTGCCTGCTCCAAAT 1186
QY 497 TCGATGAAGGAAAGCTCTCCATGACAGATCCTTCAAAGTTGGACAAAGCTTGAAC 556
Db 1187 TCGATGAAGGAAAGCTCTCCATGACAGATCCTTCAAAGTTGGACAAAGCTTGAAC 1246
QY 557 TATTAATTTATTTCCATTTCTACTGAATACGGCTAGCGGAAATTCAGAAATATGTGCAC 616
Db 1247 TATTAATTTATTTCCATTTCTACTGAATACGGCTAGCGGAAATTCAGAAATATGTGCAC 1306
QY 617 GACGAATGAATGATCTATACAAAGAAAGACTTTCGCAATTCCTCCAGATGCAGATG 676
Db 1307 GACGAATGAATGATCTATACAAAGAAAGACTTTCGCAATTCCTCCAGATGCAGATG 1366
QY 677 ACGACGACAAGCTTTACTCTCGATCGATCCAAATTTGATGAGCGAGGGAAGCTTCTTCA 736
Db 1367 ACGACGACAAGCTTTACTCTCGATCGATCCAAATTTGATGAGCGAGGGAAGCTTCTTCA 1426
QY 737 TATTTCTGTGTGATTTGACAGAGTCGAATGATATCACTAAATGCGAAAAGGAATATA 796
Db 1427 TATTTCTGTGTGATTTGACAGAGTCGAATGATATCACTAAATGCGAAAAGGAATATA 1486
QY 797 TTGAGCACACAATAATAATTTGCTCAAGCAATAATAATAATGAGAAATCCAGATATGACT 856
Db 1487 TTGAGCACACAATAATAATTTGCTCAAGCAATAATAATAATGAGAAATCCAGATATGACT 1546
QY 857 CAGACGACGTCATTTGATCAATTTAGCAAAAAGATCCATTTGATCCATGATTTGGAGAA 916
Db 1547 CAGACGACGTCATTTGATCAATTTAGCAAAAAGATCCATTTGATCCATGATTTGGAGAA 1606
QY 917 AAGTTAAGTTTGACAAAAGTTTGAGCTCATGACCCCTTGGCTGACGAATTTCAATTAAC 976
Db 1607 AAGTTAAGTTTGACAAAAGTTTGAGCTCATGACCCCTTGGCTGACGAATTTCAATTAAC 1666
QY 977 TCCAGCTGCTTCGATTTTCAAAATTTTGCAAAACCTGAGAGATCTTATTTGGGGAATGG 1036
Db 1667 TCCAGCTGCTTCGATTTTCAAAATTTTGCAAAACCTGAGAGATCTTATTTGGGGAATGG 1726
QY 1037 ATGCTCAGATGACATCTGGAAGACAGTTTCCATTCATATCAATATATACATTTATGTTGC 1096
Db 1727 ATGCTCAGATGACATCTGGAAGACAGTTTCCATTCATATCAATATATATGTTATGTTCC 1786
QY 1097 CAGTTGGTATATCGGAAAAGTATTAATTTGAACTTGTCCGCCAGATGAGTTCAAGGAA 1156

```
Db 1787 CAGTGGTATGCGGAAAGTATATATTTGAACTGTTCCGCAAGATGAGCTTCAAGGAA 1846
QY 1157 CATTGAGATGGATGATGATCTGAGAGAAATCTGAGAAACCTACCGCTTGACTGT 1216
1847 CATTGAGATGGATGATGATCTGAGAGAAATCTGAGAAACCTACCGCTTGACTGT 1906
QY 1217 TCAAGCCAAATGCCCTCCCAAGAGATAGACAAATTTAAGTAAATTCGATTTCCAAAC 1276
1907 TCAAGCCAAATGCCCTCCCAAGAGATAGACAAATTTAAGTAAATTCGATTTCCAAAC 1966
QY 1277 G----- 1277
Db 1967 GGGTGTATTATATCGTTGAGATGTTGTTCACTATTAATGATTAATGATTTCTT 2026
QY 1278 -----GCTAGAGCTACGCTTGAAAGCTGCTGACATGTTGAAATCAGTTTATTTGTC 1330
2027 GTTTTAGGTAGAGCTACGCTTGAAAGCTGCTGACATGTTGAAATCAGTTTATTTGTC 2086
QY 1331 CAGCTACAGTGAATCAGTTCAATGAGAAAGCTGATTAATTCATATTTGACGGCTGGATG 1390
2087 CAGCTACAGTGAATCAGTTCAATGAGAAAGCTGATTAATTCATATTTGACGGCTGGATG 2146
QY 1391 AAGATTTGATGAACTGATGATGATG----- 1418
2147 AAGATTTGATGAACTGATGATGATGATGATTTATCATGACCGAAGCATTTTTC 2206
Db 1419 -----CTCCATGATATTTCTACCGATGATGATGATGATGATGATGATGATGATG 1456
2207 AATGAAATTTCTATCATTTGACGCTCCATGATTTATTCACGATGATGATGATGATGATGATG 2266
QY 1457 ACAGTTATGTTCTACAACTCCGAAAGATGACAACTTTGA 1497
Db 2267 ACAGTTATGTTCTACAACTCCGAAAGATGACAACTTTGA 2307

RESULT 6
US-09-872-523-77
; Sequence 77, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 0197/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIORITY APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-77

Query Match 73.7%; Score 1103; DB 10; Length 2307;
Best Local Similarity 81.3%; Pred. No. 1,36-282;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;
```

```
QY 181 ATC-----TTGCAACA 192
111
Db 647 ATGCTGATTCATATGTTGCTAAATCGTTTAAATACAAATTTTGTAGTTGCAACA 706
QY 193 GTGCTCATGATTAATGACAAAGAACTGCGATTCGATTCAGATCGATGTTGACAGAAAT 252
707 GTGCTCATGATTAATGACAAAGAACTGCGATTCGATTCAGATCGATGTTGACAGAAAT 766
QY 253 GAAAAAGTTGCGGATACAGAGTTGCGCTCAGTTTATGCGAGCTGACAGAAATTTTGG 312
767 GAAAAAGTTGCGGATACAGAGTTGCGCTCAGTTTATGCGAGCTGACAGAAATTTTGG 826
Db 313 CTCATATTTTATCGACGATATGTTGGTTGGCAAA----- 350
827 CTCATATTTTATCGACGATATGTTGGTTGGCAAAAGATGAGCTGACGCTCTT 886
QY 351 ----- 350
887 TCTACTTCTTAATTAATATGTTCTGTCTACATAAAATTTAGAGAACATCGATTA 946
351 -----CGCCGAATGATGAT 366
947 AAACCTGCAACATTTGTATATATAGTAATAATTTGAACATTTCAAGCCGCCAATGATGAT 1006
QY 367 CCCAATATGATTAATAATTTGATATGCTCCGCCGCTTGCAATCAACGAAATACCAAAAT 426
1007 CCCAATATGATTAATAATTTGATATGCTCCGCCGCTTGCAATCAACGAAATACCAAAAT 1066
QY 427 GATATGTTAATTAATGAAAT----- 447
1067 GATATGTTAATTAATGAAATGTAAGTTTGTTCGCAATTTATGTTAATATCATCT 1126
QY 448 -----AATGCAATGATGCGCAAAATCGTGCCCAAACTGCTCTCTCCAAAT 436
1127 CACAATTCAGAAATTTGATGATGCGCAAAATCGTGCCCAAACTGCTCTCTCCAAAT 1186
QY 497 TCGATGAAAGGAAGCTCTCTTAAGCAAGATGCTTCAAGTTGGAACAGCTTGAAC 556
1187 TCGATGAAAGGAAGCTCTCTTAAGCAAGATGCTTCAAGTTGGAACAGCTTGAAC 1246
QY 557 TATTAATTTATTCGAATCTGATGAAATACGCGTAGCGCAATTCAGAAATATGTTGAC 616
1247 TATTAATTTATTCGAATCTGATGAAATACGCGTAGCGCAATTCAGAAATATGTTGAC 1306
Db 617 GACGAATGAATGATCTATACAAAGAAAGCTTCCGAATGCTTCCAGATGACATG 676
1307 GACGAATGAATGATCTATACAAAGAAAGCTTCCGAATGCTTCCAGATGACATG 1366
QY 677 ACGACAGACAAGTCTTAGCTGCGATCTCAATTTGATGAGACGAGGAAAGCTTCTCA 736
1367 ACGACAGACAAGTCTTAGCTGCGATCTCAATTTGATGAGACGAGGAAAGCTTCTCA 1426
QY 737 TATTTCTGTTGATTTGACAGCTCAATGATATCACTAAATGCGAAAAAGCAATATA 796
1427 TATTTCTGTTGATTTGACAGCTCAATGATATCACTAAATGCGAAAAAGCAATATA 1486
QY 797 TTGAGCACACAATTAATTTGCTCAGCAATTAATAATGAGAAATTCAGATATGACT 856
1487 TTGAGCACACAATTAATTTGCTCAGCAATTAATAATGAGAAATTCAGATATGACT 1546
QY 857 CAGACGACGTCACATTTGATCAATTAAGCAAAAGATCCAAATGATCCATGATTTGGGAA 916
1547 CAGACGACGTCACATTTGATCAATTAAGCAAAAGATCCAAATGATCCATGATTTGGGAA 1606
Db 917 AAGTTAAGTTTGAACAAGTTTGAAGTCAATGACCTTGCTCAGCAATTCAGATATGACT 976
1607 AAGTTAAGTTTGAACAAGTTTGAAGTCAATGACCTTGCTCAGCAATTCAGATATGACT 1666
QY 977 TCCAGCTGCTTGATTTCTCAATTTTGCAGAAAGTATCTTATTTGAGGAAATG 1036
1667 TCCAGCTGCTTGATTTCTCAATTTTGCAGAAAGTATCTTATTTGAGGAAATG 1726
Db
```


QY 1037 ATGCTCCAGATGACACTGGAAGACAGTTTCTATTCATATCATATATACATTTATGTTCC 1096
 |||||||
 Db 1727 ATGGTCCAGATGACACTGGAAGACAGTTTCTATTCATATCATATATACATTTATGTTCC 1786
 |||||||
 QY 1097 CAGTTGGTTATGCGGAAAAGTATATTTGGAACCTTGTCCGAGATGAGTTCAAGGAA 1156
 |||||||
 Db 1787 CAGTTGGTTATGCGGAAAAGTATATTTGGAACCTTGTCCGAGATGAGTTCAAGGAA 1846
 |||||||
 QY 1157 CATTCAGATGGATGAATACCTTGAGAAAAGATTCGCAAAACCCACCGCTTGACTGT 1216
 |||||||
 Db 1847 CATTCAGATGGATGAATACCTTGAGAAAAGATTCGCAAAACCCACCGCTTGACTGT 1906
 |||||||
 QY 1217 TCAGCCCAATGCTTCCCAAGAGAGATTAAGCAAAATTAAGTAAATTCGATTTCCAAAC 1276
 |||||||
 Db 1907 TCAGCCCAATGCTTCCCAAGAGAGATTAAGCAAAATTAAGTAAATTCGATTTCCAAAC 1966
 |||||||
 QY 1277 G----- 1277
 |||||||
 Db 1967 GGGTGTGTTTATATCGTTTATGAGATTTGTTTCACTATTAATAGTAAATTCATATTTGTTCTT 2026
 |||||||
 QY 1278 -----GGTAGGACTAGCGCTTGAAGCTGTGACATGTGTAATAATCGTTTATTTGTC 1330
 |||||||
 Db 2027 GTTTAAGGTAGGACTAGCGCTTGAAGCTGTGACATGTGTAATAATCGTTTATTTGTC 2086
 |||||||
 QY 1331 CAGCTACAGTGAATCAGTTTCAATGGAAGACGTATAAATGTCATTTGCGAGCTGGGATG 1390
 |||||||
 Db 2087 CAGCTACAGTGAATCAGTTTCAATGGAAGACGTATAAATGTCATTTGCGAGCTGGGATG 2146
 |||||||
 QY 1391 AAGAATTTGATGACTGTATGATGTG----- 1417
 |||||||
 Db 2147 AAGAATTTGATGACTGTATGATGTGAGTGTATATCATACCGAACGACATTTTTC 2206
 |||||||
 QY 1418 -----ACTCCATGATATTTCTACGATAGGATGGATGGTGAAGCGC 1456
 |||||||
 Db 2207 AATGAAATTTCTATCTATTTCAACTCCCATGATATTTCTACGATAGGATGGTGAAGCGC 2266
 |||||||
 QY 1457 ACAGTTATGTTCTACAACTCCGAAAAAAGTACAACTATTGA 1497
 |||||||
 Db 2267 ACAGTTATGTTCTACAACTCCGAAAAAAGTACAACTATTGA 2307
 |||||||

RESULT 7
 US-09-872-523-78
 ; Sequence 78, Application US/09872523
 ; Patent No. US20020137906A1

; GENERAL INFORMATION:
 ; APPLICANT: Horvitz, H. Robert
 ; APPLICANT: Davison, Ewa M.
 ; APPLICANT: Lu, Xiaowei
 ; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
 ; FILE REFERENCE: 01997/536002
 ; CURRENT APPLICATION NUMBER: US/09/872,523
 ; PRIOR FILING DATE: 2001-06-01
 ; PRIOR FILING DATE: 2000-06-02
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 78
 ; LENGTH: 2307
 ; TYPE: DNA
 ; ORGANISM: Caenorhabditis elegans
 US-09-872-523-78

Query Match 73.7%; Score 1103; DB 10; Length 2307;
 Best Local Similarity 81.3%; Pred. No. 1,3e-282;
 Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;

QY 1 ATGCTGAATTTCTGAAAAATGTCAGAGCTACAAAAAATCGACAGAAAACCTCGATAG 60
 |||||||
 Db 467 ATGCTGAATTTCTGAAAAATGTCAGAGCTACAAAAAATCGACAGAAAACCTCGATAG 526
 |||||||
 QY 61 ACCTACTGTGGAAATCCATTTACATCAGTTGGAAGAAAACCTCTTTCATTTCCA 120
 |||||||

Db 527 ACCTACTGTGGAAATCCATTTACATCAGTTGGAAGAAAACCTCTTTCATTTCCA 586
 |||||||
 QY 121 GTTGAAGCATTCGAATCGTAACCTTAACGTAAATTTAAAGATGGGTGAAGAGAGTT 180
 |||||||
 Db 587 GTTGAAGCATTCGAATCGTAACCTTAACGTAAATTTAAAGATGGGTGAAGAGAGTT 646
 |||||||
 QY 181 ATC-----TTGGAACA 192
 |||||||
 Db 647 ATCGAGAGTTCAATATTTGTTCTGTAATTCGTTTAAATTCATATTTTGTAGTTCCAAACA 706
 |||||||
 QY 193 GTGCTCATGATTAATGACAAAGACCTGCATTTGATTTCAAGTCAGATGTTTGCAGCAAT 252
 |||||||
 Db 707 GTGCTCATGATTAATGACAAAGACCTGCATTTGATTTCAAGTCAGATGTTTGCAGCAAT 766
 |||||||
 QY 253 GAAAAAGTTGGGAGATACAGATTCGCTGCTGATTAATGAGAGTCAGACGAAATTTGG 312
 |||||||
 Db 767 GAAAAAGTTGGGAGATACAGATTCGCTGCTGATTAATGAGAGTCAGACGAAATTTGG 826
 |||||||
 QY 313 CTCATATATTTATCCGACGATATGTTGGTTGGCAAA----- 350
 |||||||
 Db 827 CTCATATATTTATCCGACGATATGTTGGTTGGCAAAATAGTTGACGTCAGCTGTT 886
 |||||||
 QY 351 ----- 350
 |||||||
 Db 887 TCTACTATTTCTAAATTAATTAATGTTCTGTTTACATTAATAATTCAGAGAAATCGTATTA 946
 |||||||
 QY 351 -----CGCCGCAATGAGTAT 366
 |||||||
 Db 947 AAATCTCGAAACATTTGTATTAATAGTAAATTTGAACATTTTACAGCCGCCCAATGAGTAT 1006
 |||||||
 QY 367 CCCAATATGATTAATAATGTAATATGCTCCGCGCTTGCATACAGAAATACCAAAAT 426
 |||||||
 Db 1007 CCCAATATGATTAATAATGTAATATGCTCCGCGCTTGCATACAGAAATACCAAAAT 1066
 |||||||
 QY 427 GATATGTAATTAATTAAT----- 447
 |||||||
 Db 1067 GATATGTAATTAATTAATTAATGTTTTCGCAATTAATGTAATTAATCATCT 1126
 |||||||
 QY 448 -----AATTCATTTGATGGCAATGCGGCAAACTTGCCTGCTGCCAAAT 496
 |||||||
 Db 1127 CACAACCTTAATAATTTGATTTGATGCGGAAATGTCGCGCAAACTTGCCTGCTGCCAAAT 1186
 |||||||
 QY 497 TCGATGAAGGAAGGCTCTCTTAAGCAACATGTTTCAAGTTGGACACGCTTTGAAC 556
 |||||||
 Db 1187 TCGATGAAGGAAGGCTCTCTTAAGCAACATGTTTCAAGTTGGACACGCTTTGAAC 1246
 |||||||
 QY 557 TATTAATTAATTTCAATTTCTACTGAATACGCGTACGCGGAATTCAGAAATATGTGAC 616
 |||||||
 Db 1247 TATTAATTAATTTCAATTTCTACTGAATACGCGTACGCGGAATTCAGAAATATGTGAC 1306
 |||||||
 QY 617 GACGATGAATGATCTATCACAAAGAAAGACCTTCCGAATCGCTTCAGATGCGAGATG 676
 |||||||
 Db 1307 GACGATGAATGATCTATCACAAAGAAAGACCTTCCGAATCGCTTCAGATGCGAGATG 1366
 |||||||
 QY 677 ACGACAGACAGTCTTTAGCTCTGATCTCAATATTTGATAGACGAGGAAAGCTTTCTCA 736
 |||||||
 Db 1367 ACGACAGACAGTCTTTAGCTCTGATCTCAATATTTGATAGACGAGGAAAGCTTTCTCA 1426
 |||||||
 QY 737 TATTTCTGTTGATTTGACAGATCAATGATATCACTAAATGCGAAAAAGGAATATA 796
 |||||||
 Db 1427 TATTTCTGTTGATTTGACAGATCAATGATATCACTAAATGCGAAAAAGGAATATA 1486
 |||||||
 QY 797 TTGACGACGCAATTAATAATGTCGAAGCAATTAATAATGCGAAAAATCCAAATATGATC 856
 |||||||
 Db 1487 TTGACGACGCAATTAATAATGTCGAAGCAATTAATAATGCGAAAAATCCAAATATGATC 1546
 |||||||
 QY 857 CAGACAGCTCATATTTGATCAATTAAGCAAAAGATTCATTAATGCCATGATTTGGAGAA 916
 |||||||
 Db 1547 CAGACAGCTCATATTTGATCAATTAAGCAAAAGATTCATTAATGCCATGATTTGGAGAA 1606
 |||||||
 QY 917 AAGTTAAGTTGGACAAAAGTTTGAAGCTCATGACCCCTTGGCTCAGCAATTCATTAAC 976
 |||||||
 Db 1607 AAGTTAAGTTGGACAAAAGTTTGAAGCTCATGACCCCTTGGCTCAGCAATTCATTAAC 1666
 |||||||

```
OY 977 TCACGCTGCTTCATTCCTCAAAATTTTGCAAAAGAGATATCTATTGTGGGAATG 1036
|||||
1667 TCACGCTGCTTCATTCCTCAAAATTTTGCAAAAGAGATATCTATTGTGGGAATG 1726
OY 1037 ATGCTCCAGATGCAGCTTGAAGACAGTTTCTCATTCATATCAATATCATATTATGTTCC 1096
|||||
Db 1727 ATGCTCCAGATGCAGCTTGAAGACAGTTTCTCATTCATATCAATATCATATTATGTTCC 1786
OY 1097 CAGTTGGTTATGCGGAAAGATATATTTGGAACCTGTTCCGCGAGATGAGTTCAAGGAA 1156
|||||
Db 1787 CAGTTGGTTATGCGGAAAGATATATTTGGAACCTGTTCCGCGAGATGAGTTCAAGGAA 1846
OY 1157 CATTCAAGTGGAGTAAATCTTGGAGAAAGATCTGCAAGAAACCTTACCCTTACCTTGT 1216
|||||
Db 1847 CATTCAAGTGGAGTAAATCTTGGAGAAAGATCTGCAAGAAACCTTACCCTTACCTTGT 1906
OY 1217 TCACGCCAATGCTTCCCAAGAGATTTAGACAAATTTAAGTATATCTGATTTCCAAAC 1276
|||||
Db 1907 TCACGCCAATGCTTCCCAAGAGATTTAGACAAATTTAAGTATATCTGATTTCCAAAC 1966
OY 1277 G----- 1277
1967 GGGTGTGTTTATCGTTTGATGTTTTCACATTAATAGTATATCATATTTGTTCTT 2026
OY 1278 -----GTAAGACATACGCTTGAAGCTGCTGACATGCTGAAATCAGTTTATTTGTC 1330
|||||
Db 2027 GTTTTAAGGTAGGACTACGCTTGAAGCTGCTGACATGCTGAAATCAGTTTATTTGTC 2086
OY 1331 CAGCTACAGTGAATCAGTTTCAATGATGATGATGATGATGATGATGATGATGATGATG 1390
|||||
Db 2087 CAGCTACAGTGAATCAGTTTCAATGATGATGATGATGATGATGATGATGATGATGATG 2146
OY 1391 AAGAATTTGATGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1418
|||||
Db 2147 AAGAATTTGATGAACCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2206
OY 1419 -----CTCCCATGATATTTCTACCGATGATGATGATGATGATGATGATGATG 1456
|||||
Db 2207 AATGAAATTTCTATTCATTCAGCTCCATGATGATGATGATGATGATGATGATGATG 2266
OY 1457 ACAGTTATGTTCTACAACTCCGAAAGAGTACAACTATTTGA 1497
|||||
Db 2267 ACAGTTATGTTCTACAACTCCGAAAGAGTACAACTATTTGA 2307
```

RESULT 8
US-09-801-574-11
; Sequence 11, Application US/09801574
; Patent No. US20020081592A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Beijing Jeremy
; TITLE OF INVENTION: Reproduction-Specific Genes
; FILE REFERENCE: 0399, 2007-002
; CURRENT APPLICATION NUMBER: US/09/801,574
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,518
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/261,557
; PRIOR FILING DATE: 2001-01-12
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 2969
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-801-574-11

Query Match 3.9%; Score 58.8; DB 9; Length 2969;
Best Local Similarity 55.3%; Pred. No. 6.6e-05;
Matches 114; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

```
OY 1279 GTAGACTACGCCCTTGAAGCTGCTGACATGTGTGAATAATAGTTATTGTCACGCTACA 1338
|||||
Db 801 GTTGGGATGAGAGATGAGGCTGTGATGAGAAAAATCCATTATGATCTGTCGACACA 860
OY 1339 GTGAATCAGTTCATGGAAGACATGATTAATGTCAATTTTCGACGCTGGATGAAGAAATTT 1398
|||||
Db 861 ATTGAGCTGTCTGTGTGAGATCAACTCATATCTCTTTGATGATGATGATGATGATGATGAT 920
OY 1399 GATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
|||||
Db 921 GATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 980
OY 1459 AGTTATGTTCTACAACTCCGAAAAA 1484
|||||
Db 981 GGAGATGATTAATACAGCCACAGGAAA 1006
```

RESULT 9
US-09-789-919-55
; Sequence 55, Application US/09789919
; Patent No. US20020064855A1
; GENERAL INFORMATION:
; APPLICANT: Lemischka, Ihor
; TITLE OF INVENTION: GENES THAT REGULATE HEMATOPOIETIC BLOOD FORMING STEM
; FILE REFERENCE: 2275-1-005
; CURRENT APPLICATION NUMBER: US/09/789,919
; CURRENT FILING DATE: 2001-02-21
; NUMBER OF SEQ ID NOS: 96
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 55
; LENGTH: 3590
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-789-919-55

Query Match 3.8%; Score 57.4; DB 9; Length 3590;
Best Local Similarity 55.2%; Pred. No. 0.00017;
Matches 112; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

```
OY 1276 CGGCTAGGACTACGCCCTTGAAGCTGCTGACATGTGTGAATAATCAGTTTATTTGTCAGCT 1335
|||||
Db 1745 CGGTTAGGAATGAATTAAGAGCTGTAGATCATGAGCCAGGTTAATATGTTGAGCC 1804
OY 1336 ACAGTGAATCAGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1395
|||||
Db 1805 ACAGTGAATCAGTTTCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1864
OY 1396 TTTGATGACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455
|||||
Db 1865 TATGACCGTGGTGAAGCTGTGATGATGATGATGATGATGATGATGATGATGATGATG 1924
OY 1456 CACAGTTATGTTCTACAACTCC 1478
|||||
Db 1925 ACTGATATCAACTACAGCTCC 1947
```

RESULT 10
US-09-764-868-396
; Sequence 396, Application US/09764868
; Patent No. US2002016871A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.,
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PAM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 396
; LENGTH: 2403


```
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (1)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (10)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (18)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (41)
OTHER INFORMATION: n equals a,t,g, or c
NAME/KEY: SITE
LOCATION: (2398)
OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-396
```

Query Match 3.7%; Score 55.2; DB 10; Length 2403;

Best Local Similarity 55.9%; Pred. No. 0.00053; Mismatches 83; Indels 0; Gaps 0;

Matches 105; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

```
QY 1291 CTTGAAGCTGCTGACATGTGTGAAATCAGTTATTGTCACAGTACAGTGAATCAGTT 1350
DB 700 CTTGAGGTGTGACAAAAGAACCCCTATGTTATAGAGTACAGTGTGGACAGACA 759
QY 1351 CATGAAGACTGTAAATGTCATTTTCGACGGCTGGATGAGAAATTTGATGACTGTAT 1410
DB 760 GATGATCACCAGGTTAAAGTTCACCTTGATGCTGGACAAATGCTATGATTCCTGATA 819
QY 1411 GATGTGACCTCCATGATTTCTACGATAGATGATGATGATGATGATGATGATGATGAT 1470
DB 820 GATGACGATTTCTCTGATATTCACCTGTAGGCTGTGTTCAAAAACAGACATCCCTT 879
QY 1471 CAACCTCC 1478
DB 880 CAGCTCC 887
```

RESULT 11

US-10-204-887-14

Sequence 14, Application US/10204887

Publication No. US20030124569A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: PANZER, Scott R.

APPLICANT: SPIRO, Peter A.

APPLICANT: BANVILLE, Steven C.

APPLICANT: SHAH, Purvi

APPLICANT: CHATLUP, Michael S.

APPLICANT: CHAN, Simon C.

APPLICANT: CHEN, Alice

APPLICANT: D/SA, Steven A.

APPLICANT: AMSHEY, Stefan

APPLICANT: DAHL, Christopher R.

APPLICANT: DAM, Tam C.

APPLICANT: DANIELS, Susan E.

APPLICANT: DUFOUR, Gerard E.

APPLICANT: FLORES, Vincent

APPLICANT: FONG, Willy T.

APPLICANT: GREENAWALT, Lila B.

APPLICANT: HILLMAN, Jennifer L.

APPLICANT: JONES, Anissa L.

APPLICANT: LIU, Tommy F.

APPLICANT: ROSEBERRY, Ann M.

APPLICANT: ROSEN, Bruce H.

APPLICANT: RUSSO, Frank D.

APPLICANT: STOCKDREHER, Theresa K.

APPLICANT: DAFEO, Abel

APPLICANT: WRIGHT, Rachel J.

APPLICANT: YAP, Pierre E.

```
APPLICANT: YU, Jimmy Y.
APPLICANT: BRADLEY, Diana L.
APPLICANT: BRATCHER, Shawn R.
APPLICANT: CHEN, Wensheng
APPLICANT: COHEN, Howard J.
APPLICANT: HODGSON, David M.
APPLICANT: LINCOLN, Stephen E.
TITLE OF INVENTION: SECRETORY MOLECULES
FILE REFERENCE: PT-1134 PCT
CURRENT APPLICATION NUMBER: US/10/204,887
CURRENT FILING DATE: 2002-08-21
PRIOR APPLICATION NUMBER: 60/185,215; 60/185,216; 60/205,232; 60/205,323; 60/205,2
60/205,324; 60/205,286
PRIOR FILING DATE: 2000-02-24; 2000-02-24; 2000-05-16; 2000-05-17; 2000-05-17;
2000-05-17; 2000-05-17
NUMBER OF SEQ ID NOS: 159
SOFTWARE: PERL Program
SEQ ID NO 14
LENGTH: 1035
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. US20030124569A1 LI:228319.1:2000PEB01
US-10-204-887-14
```

Query Match 3.6%; Score 53.4; DB 14; Length 1035;

Best Local Similarity 55.0%; Pred. No. 0.00099; Mismatches 86; Indels 0; Gaps 0;

Matches 105; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

```
QY 1291 CTTGAAGCTGCTGACATGTGTGAAATCAGTTATTGTCACAGTACAGTGAATCAGTT 1350
DB 815 CTTGAAGTGTGTGATTAACGGAACCCAGTTATTCGTTGCTACGATTTGATGATTT 874
QY 1351 CATGAAGACTGTAAATGTCATTTTCGACGGCTGGATGAGAAATTTGATGACTGTAT 1410
DB 875 GATGACCAAGATTAAGTTTCATTTGATGTTGGACATTAATGATGACTACTGGTG 934
QY 1411 GATGTGACCTCCATGATTTCTACGATAGATGATGATGATGATGATGATGATGATGAT 1470
DB 935 GAGGACAGACGCCCTGATATTCACCCGATGCGATTTGTGATGTCACAGGGCATTCACATG 994
QY 1471 CAACCTCCGAA 1481
DB 995 GAAGTCCCAA 1005
```

RESULT 12

US-09-864-761-20410

Sequence 20410, Application US/09864761

Patent No. US20020048763A1

GENERAL INFORMATION:

APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.

APPLICANT: Hanzel, David K.

APPLICANT: Chen, Wensheng

TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO

FILE REFERENCE: Aemica-X-1

CURRENT APPLICATION NUMBER: US/09/864,761

CURRENT FILING DATE: 2001-05-23

PRIOR APPLICATION NUMBER: US 60/180,312

PRIOR FILING DATE: 2000-02-04

PRIOR APPLICATION NUMBER: US 60/207,456

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: US 09/632,366

PRIOR FILING DATE: 2000-08-03

PRIOR APPLICATION NUMBER: GB 24263.6

PRIOR FILING DATE: 2000-10-04

PRIOR APPLICATION NUMBER: US 60/236,359

PRIOR FILING DATE: 2000-09-27

PRIOR APPLICATION NUMBER: PCT/US01/00666

PRIOR FILING DATE: 2001-01-30

```
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 20410
;; LENGTH: 204
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL110502.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.93
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;; OTHER INFORMATION: SWISSPROT HIT: Q01441, EVALUE 2.20e+00
;; OTHER INFORMATION: EST HUMAN HIT: BE256221.1, EVALUE 1.00e-111
;; OTHER INFORMATION: NT HIT: g16912641, EVALUE 1.00e-111
US-09-864-761-20410

Query Match          3.5%; Score 52; DB 9; Length 204;
Best Local Similarity 54.8%; Pred. No. 0.0092;
Matches 103; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

QY 1291 CTTAAGCTGCTGACATGTGTGAATCAGTTTATTGTCCAGCTACAGTAAATCACTT 1350
    || || || || || || || || || || || || || || || || || || || || ||
DB 6 CTAGAAGCTGTGACAGAGAGAACACCTCATTTTCATTTGCCAGCACAATTTGGGAGGCTT 65
QY 1351 CATGGAAGACATGATAAATGTCATTTGAGGCGTGGGATGGAAGATTGATGAACTGTAT 1410
    || || || || || || || || || || || || || || || || || || || || ||
DB 66 CGGGGCTCAGAGGTGCTTGTACTTGTGATGGTGGCGAGGGGCTTGTGACTACTGTGTC 125
QY 1411 GATGTGACATCCCATGATATTTCTACCGATAGATGATGTGAAGGCGACAGTTTATGTTCTA 1470
    || || || || || || || || || || || || || || || || || || || || ||
DB 126 CGCTTCGACCTCCGAGACATCTTCCCTGTGGCTGTGTCTCTTGTGACTGAGACAACTGTG 185
QY 1471 CAACCTCC 1478
    || || || || ||
DB 186 CAGCTCC 193

RESULT 13
US-09-864-761-3642
; Sequence 3642, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeonica-X-1
```

```
;; CURRENT APPLICATION NUMBER: US/09/864,761
;; CURRENT FILING DATE: 2001-05-23
;; PRIOR APPLICATION NUMBER: US 60/180,312
;; PRIOR FILING DATE: 2000-02-04
;; PRIOR APPLICATION NUMBER: US 60/207,456
;; PRIOR FILING DATE: 2000-05-26
;; PRIOR APPLICATION NUMBER: US 09/632,366
;; PRIOR FILING DATE: 2000-08-03
;; PRIOR APPLICATION NUMBER: GB 24263,6
;; PRIOR FILING DATE: 2000-10-04
;; PRIOR APPLICATION NUMBER: US 60/236,359
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: PCT/US01/00666
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 3642
;; LENGTH: 494
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AL110502.1
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.93
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.1
US-09-864-761-3642

Query Match          3.4%; Score 51.6; DB 9; Length 494;
Best Local Similarity 56.5%; Pred. No. 0.0019;
Matches 96; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 1280 TAGACATGCGCTTGAAGCTCTGACATGTGTGAATATCACTTATTTGTCACACTCAG 1339
    || || || || || || || || || || || || || || || || || || || || ||
DB 300 TGGGAATGAAGCTAGAACCTGTGACAGAGAAGAACCCCTCATTTATTTGGCCACCACTA 359
QY 1340 TGAATACGCTCATGAGAACTGATTAATGTCAATTTGACAGCGCTGGGATGAAGAAATTTG 1399
    || || || || || || || || || || || || || || || || || || || || ||
DB 360 TTGGGAGAGTTCGGGGCTCAGAGAGTGTCTGTGACTTTTGTGATGGGTGGCGAGGGCTTTTG 419
QY 1400 ATGAACGTATGATGTGACATCCCATGATATTTCAACCGATAGATGTGT 1449
    || || || || || || || || || || || || || || || || || || || || ||
DB 420 ACTACTGGTCCGCTTGACATCCGAGACATCTTCCCTGTGGCTGTGT 469

RESULT 14
US-09-918-995-35932
; Sequence 35932, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
```

```

1  APPLICANT: Hyseq, Inc.
2  TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
3  TITLE OF INVENTION: FROM VARIOUS CDNA LIBRARIES
4  FILE REFERENCE: 20411-756
5  CURRENT APPLICATION NUMBER: US/09/918,995
6  CURRENT FILING DATE: 2001-07-30
7  PRIOR APPLICATION NUMBER: US/09/235,076
8  PRIOR FILING DATE: 1999-01-20
9  NUMBER OF SEQ ID NOS: 38054
10 SOFTWARE: FASTSEQ for Windows Version 3.0
11 SEQ ID NO 35932
12 LENGTH: 413
13 TYPE: DNA
14 ORGANISM: Homo sapiens
15 US-09-918-995-35932

```

Query Match	3.48;	Score 51;	DB 11;	Length 413;
Best Local Similarity	54.58;	Pred. NO. 0.0025;		
Matches 102;	Conservative	0;	Mismatches 85;	Indels 0;
				Gaps 0;

QY	1292	TTTGAAGTGGCGCATGTGTTAAATCATGTTTATTTTCACAGTACAGTGAATCAAGTTC	1355
Dd	5	TAGAAGCAGTAGATCTCATGAGCCACGCTTTAAATATGTGTAGCCACAGTAACCTGCAATTA	64
QY	1352	ATGGAAGCATGATAATGTCAATTTTCAGCGCTGGATGAAGAAATTTGATGAACGTATG	1411
Dd	65	TTTCATCGCTCTTAGGATACATTTTGAATGGATGGGAAGAAGATATCATCTAGTGGTAG	124
QY	1412	ATGTGGACCTCCATGATATTCTACCGATAGAGATGGTGAAGCGCACAGTTATGTTTAC	1471
Dd	125	ACTGTGAGTCACTGACCTATCTATCCTTAGGAGGTGGTCACTTAACGTGATATCAACTAC	184
QY	1472	AACCTCC	1478
Dd	185	AGCCTCC	191

```

RESULT 15
US-10-027-632-103678
: Sequence 103678, Application US/10027632
: GENERAL INFORMATION:
: APPLICANT: Wang, David G.
: TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
: POLYMORPHISMS IN THE HUMAN GENOME
: FILE REFERENCE: 108827.129
: CURRENT APPLICATION NUMBER: US/10/027,632
: CURRENT FILING DATE: 2002-04-30
: PRIOR APPLICATION NUMBER: US 60/219,006
: PRIOR FILING DATE: 2000-07-12
: PRIOR APPLICATION NUMBER: US 60/198,676
: PRIOR FILING DATE: 2000-04-20
: PRIOR APPLICATION NUMBER: US 60/193,483
: PRIOR FILING DATE: 2000-03-29
: PRIOR APPLICATION NUMBER: US 60/185,218
: PRIOR FILING DATE: 2000-02-24
: PRIOR APPLICATION NUMBER: US 60/167,363
: PRIOR FILING DATE: 1999-11-23
: PRIOR APPLICATION NUMBER: US 60/156,358
: PRIOR FILING DATE: 1999-09-28
: PRIOR APPLICATION NUMBER: US 60/146,002
: PRIOR FILING DATE: 1999-08-09
: NUMBER OF SEQ ID NOS: 325720
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 103678
: LENGTH: 2445
: TYPE: DNA
: ORGANISM: Human
: US-10-027-632-103678

```

Query Match	3.4%	Score 50.6;	DB 13;	Length 2445;
Best Local Similarity	53.2%;	Pred. No. 0.009;		
Matches 107; Conservative	0;	Mismatches 94;	Indels 0;	Gaps 0;

Oy	1278	GGTGGAGCTACGCGCTTGAAGCTGGTGACATGTGAAATTAAGTTTATTGTCACGGTAC	1337
Dd	1412	GGTGGGCCTGAAGCTGGAGGCCGTGGACCTTATGTAGGCCCGCGCTCATCTGTGGCCAC	1471
Oy	1338	AGTGAATTCATCTCATGTGGAGACTGATMAATGTCAATTTGACGCGCTGGATGGAAGATT	1397
Dd	1472	GGTGAACGAGTGGTGGCATCGGCTCCTCAGCATCCATCTTAGCGGTGGACACGAGATA	1531
Oy	1398	TGATGAACCTGATGATGTGACCTCCCATGATATTTACCGATAGAGATGCTGAAACGCA	1457
Dd	1532	CGACCAGGGGTGAGACTGCGAGTCCCCACAGACATCATCCCCGTGGCTGTGAGCTCAC	1591
Oy	1458	CAGTTATGTTTACAACCTCC	1478
Dd	1592	CGGCTACCAAGCTCAGCCTCC	1612

Search completed: September 2, 2003, 17:50:50
Job time : 419 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 14:53:19 ; Search time 3439 Seconds
(without alignments)
10579.762 Million cell updates/sec

Title: US-09-872-523-6
Perfect score: 1497
Sequence: 1 atgtctgaattcttgaat.....cgaaaagracactatga 1497

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estov:*
6: em_estopl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estrom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vr1:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	630.4	42.1	9 AU204939	AU204939 AU204939
2	559	40.0	768 12 BU143186	BU143186 BU143186
3	591	39.5	761 12 BU129361	BU129361 BU129361
4	563.4	37.6	607 14 CB404801	CB404801 OSTR028D3

5	559	37.3	581 12 BU101701	BU101701 BU101701
6	521	34.8	557 12 BU124772	BU124772 BU124772
7	519	34.7	677 12 AU216908	AU216908 AU216908
8	510	34.1	686 12 BU153087	BU153087 BU153087
9	502.4	33.6	519 12 BU115564	BU115564 BU115564
10	366	24.4	380 9 AV197399	AV197399 AV197399
11	360	24.0	360 14 CA42305	CA42305 CA42305
12	308	20.6	340 9 AU111588	AU111588 AU111588
13	220	14.7	300 9 AU116296	AU116296 AU116296
14	215	14.4	300 9 AU115428	AU115428 AU115428
15	186.2	12.4	337 9 AU112450	AU112450 AU112450
16	80.2	5.4	377 14 CA3257	CA3257 CA3257
17	66	4.4	505 14 CB385690	CB385690 OSTR028D3
18	65	4.3	360 14 C70202	C70202 C70202
19	60.6	4.0	554 10 BE684409	BE684409 184807 MA
20	59.8	4.0	291 14 CB276655	CB276655 PL34P01.Y
21	59	3.9	350 9 AAB74037	AAB74037 v87e01.F
22	59	3.9	521 13 BQ552847	BQ552847 H4018E12-
23	59	3.9	548 9 AU045245	AU045245 AU045245
24	59	3.9	605 28 A2952671	A2952671 2M0217P03
25	59	3.9	655 10 BB068384	BB068384 BB068384
26	59	3.9	664 13 BQ552848	BQ552848 H4018E12-
27	59	3.9	4114 11 AK028503	AK028503 Mus muscu
28	59	3.9	4852 11 AK033238	AK033238 Mus muscu
29	58.8	3.9	2938 11 AK016533	AK016533 Mus muscu
30	58.8	3.9	3442 11 AK040521	AK040521 Mus muscu
31	58.8	3.9	4845 11 AK031490	AK031490 Mus muscu
32	58.6	3.9	346 10 BG057154	BG057154 UI-R-G00-
33	57.4	3.8	724 14 CD373179	CD373179 UI-R-G00-
34	57.4	3.8	3590 13 BU699337	BU699337 sa49P01-F
35	56.6	3.8	784 13 BY395971	BY395971 BY395971
36	56.6	3.8	784 13 CA510759	CA510759 UI-R-FJ0-
37	56.2	3.8	630 14 CB57560	CB57560 AMGNMUC:N
38	55.8	3.7	408 9 AL673391	AL673391 AL673391
39	55.8	3.7	476 10 BE376727	BE376727 RC6-PN007
40	55.8	3.7	593 10 BG348697	BG348697 da016F01.
41	55.8	3.7	631 9 AL640068	AL640068 AL640068
42	55.8	3.7	804 9 AL879839	AL879839 AL879839
43	55.8	3.7	2048 11 BC034364	BC034364 Homo sap1
44	55.4	3.7	603 14 CB581904	CB581904 AMGNMUC:N
45	55	3.7	379 9 AM489725	AM489725 UT-M-BH3-

ALIGNMENTS

RESULT 1
AU204939 653 bp mRNA linear EST 17-JUL-2001
LOCUS AU204939 unpublished oligo-capped cDNA library, stage I4
DEFINITION Caenorhabditis elegans cDNA clone yk843e05 5', mRNA sequence.
ACCESSION AU204939 GI:14836792
VERSION AU204939.1
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peloderiinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 653)
AUTHORS Kohara,Y., Shin-I,T., Thlerry-Mleg,J., Thlerry-Mleg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 653
/organism="Caenorhabditis elegans"

FEATURES
source

```

/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk843e05"
/sex="Hermaphrodite"
/rissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cDNA library, stage
L4"

BASE COUNT      216 a      116 c      136 g      185 t
ORIGIN
Query Match      42.1%; Score 630.4; DB 9; Length 653;
Best Local Similarity 99.8%; Pred. No. 1.6e-116;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCTGGAATTTGCAAAATGTCAGCTAACAAAAATCGACAGAAACTCGATAG 60
    |||||||
DB 22 ATGCTGGAATTTGCAAAATGTCAGCTAACAAAAATCGACAGAAACTCGATAG 81
    |||||||
OY 61 ACCTACTGTGGGAATCCTATTTCATCAGTTTCGAGAAAGAAAACTTTTCATTC 120
    |||||||
DB 82 ACCTACTGTGGGAATCCTATTTCATCAGTTTCGAGAAAGAAAACTTTTCATTC 141
    |||||||
OY 121 GTTGAAGCATTCACATCGTAACCTTAACAGTAATTTTAAAGATGCGTGAAGAGAG 180
    |||||||
DB 142 GTTGAAGCATTCACATCGTAACCTTAACAGTAATTTTAAAGATGCGTGAAGAGAG 201
    |||||||
OY 181 ATCTTCGAAAGAGTGTCCATGATTATGACAAAGAACTCGATTCGATTCAGATCAGAT 240
    |||||||
DB 202 ATCTTCGAAAGAGTGTCCATGATTATGACAAAGAACTCGATTCGATTCAGATCAGAT 261
    |||||||
OY 241 TTTGCAGCAATTTGAAAAAGTTTGGGAGATACAGAGTTCTGCTCAAGTTTATCGAGCTGAC 300
    |||||||
DB 262 TTTGCAGCAATTTGAAAAAGTTTGGGAGATACAGAGTTCTGCTCAAGTTTATCGAGCTGAC 321
    |||||||
OY 301 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGGTTTGGCAACGCCGCAATG 360
    |||||||
DB 322 ACGAAATTTTGGCTCAATATTTTATCGAGCATATGTTGGTTTGGCAACGCCGCAATG 381
    |||||||
OY 361 AGTGATCCCATATGATGATTAATGATATGCTCGCGGCTTGCATTAACGAGAAATAG 420
    |||||||
DB 382 AGTGATCCCATATGATGATTAATGATATGCTCGCGGCTTGCATTAACGAGAAATAG 441
    |||||||
OY 421 CAAATGATATGTAATTAATTAATTAATGATGATGAGGAAATGCTGCGCCAAACT 480
    |||||||
DB 442 CAAATGATATGTAATTAATTAATTAATGATGATGAGGAAATGCTGCGCCAAACT 501
    |||||||
OY 481 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTTAAGCAAGCATCGTTTCAAAGTT 540
    |||||||
DB 502 TCGCTGTCTCCAAAATTTGATGAAGGAGGCTCTCTTAAGCAAGCATCGTTTCAAAGTT 561
    |||||||
OY 541 GGACAAAGCTTTGAACCTTTAAATTAATTCATTAATTCAGGAGTGAAGGAGGAGT 600
    |||||||
DB 562 GGACAAAGCTTTGAACCTTTAAATTAATTCATTAATTCAGGAGTGAAGGAGGAGT 621
    |||||||
OY 601 CAAGAAATATGTGAAGCAAGATGAATGATATC 632
    |||||||
DB 622 CAAGAAATATGTGAAGCAAGATGAATGATATC 653
    |||||||

RESULT 2
BU143186/ 768 bp mRNA linear EST 23-Jan-2002
LOCUS BU143186 unpublished oligo-capped cDNA library, C. elegans L4 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1192h09 3', mRNA sequence.
ACCESSION BU143186
VERSION BU143186.1 GI:18303352
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae
; Rhabditidae; Pelodierinae; Caenorhabditis.

```

```

REFERENCE 1 (bases 1 to 768)
AUTHORS Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-i
National Institute of Genetics
111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
source
1. 768
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1192h09"
/sex="hermaphrodite"
/rissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L4 stage"

BASE COUNT      232 a      151 c      146 g      237 t      2 others
ORIGIN
Query Match      40.0%; Score 599; DB 12; Length 768;
Best Local Similarity 96.8%; Pred. No. 3.4e-110;
Matches 630; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

OY 847 AGATATGACTCAGACGACGTCACATTTGATGATTAAGCAAAAGATCCATATGATCCATG 906
    |||||||
DB 768 AGATATGACTCAGACGACGTCACATTTGATGATTAAGCAAAAGATCCATATGATCCATG 709
    |||||||
OY 907 ATTGGAGAAAGTTAAGTTGGACAAAGTTTACGTCATGCAAGCCCTGGCTCAGCA 966
    |||||||
DB 708 ATTGGAGAAAGTTAAGTTGGACAAAGTTTACGTCATGCAAGCCCTGGCTCAGCA 649
    |||||||
OY 967 TTCAATAACCTCCACGTCGCTTCGATTCCTCAATTTTGGCAAAAGTGAAGATATCTTAT 1026
    |||||||
DB 648 TTCAATAACCTCCACGTCGCTTCGATTCCTCAATTTTGGCAAAAGTGAAGATATCTTAT 589
    |||||||
OY 1027 GTGGGAATGATGTCAGATGACATCTTGAGACAGCTTTCTTATTCATATCAATTAATCA 1086
    |||||||
DB 588 GTGGGAATGATGTCAGATGACATCTTGAGACAGCTTTCTTATTCATATCAATTAATCA 529
    |||||||
OY 1087 TTTATGCTCCAGTTGGTTATGCGGAAAGATATTAATTTGAACTTGTCCGCAAGATGAG 1146
    |||||||
DB 528 TTTATGCTCCAGTTGGTTATGCGGAAAGATATTAATTTGAACTTGTTCGCGCAAGATGAG 469
    |||||||
OY 1147 TTCAAGGACATTCAGATGAGATGAATTAATTTGAGAAAGATCTGCAAAACCCCTACCG 1206
    |||||||
DB 468 TTCAAGGACATTCAGATGAGATGAATTAATTTGAGAAAGATCTGCAAAACCCCTACCG 409
    |||||||
OY 1207 CTGACTTGTTCAAAGCCCATGCTTTCCCAAGAGATATGACAAATTTAAGCTAATTCG 1266
    |||||||
DB 408 CTGACTTGTTCAAAGCCCATGCTTTCCCAAGAGATATGACAAATTTAAGCTAATTCG 359
    |||||||
OY 1267 ATTTCCAAAGGGGTAGGACTAGGCTGAGCTGTCATGTCATGTCATGTCATGTCATGTCAT 1326
    |||||||
DB 358 -----GTRGGACTACGCTTGAACCTGTCATGTCATGTCATGTCATGTCATGTCAT 310
    |||||||
OY 1327 TGTCAGCTACAGTGAATCAATTCATGAGAGCTGATTAATGTCATTTTCAGAGGCTGAG 1386
    |||||||
DB 309 TGTCAGCTACAGTGAATCAATTCATGAGAGCTGATTAATGTCATTTTCAGAGGCTGAG 250
    |||||||
OY 1387 GATGAAGATTTGATGAACCTGTATGATGTCAGCTGCCATGATATTCACGATATGATGAG 1446
    |||||||
DB 249 GATGAAGATTTGATGAACCTGTATGATGTCAGCTGCCATGATATTCACGATATGATGAG 190
    |||||||
OY 1447 TGTCAGGCGCAGTATGTCATCAACCTCGGAAAAAGTCAACTATATGA 1497
    |||||||

```

```

Db      189 TGTAAGCGCAGATTATGTTCTACAACTCCGAAAAAGTACATTATGA 139

RESULT 3
LOCUS   B129361 761 bp mRNA linear EST 23-JAN-2002
DEFINITION B129361 unpublished oligo-capped cDNA library, C. elegans L1 stage
ACCESSION B129361
VERSION   B129361.1 GI:18289518
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
          ; Rhabditidae; Peleoderinae; Caenorhabditis.
REFERENCE Kohara,Y., Shin-I.T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
          and Sugano,S.
          A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT  Contact: Tadasi Shin-I
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshin@genes.nig.ac.jp.
          Location/Qualifiers
FEATURES
    source             1..761
                        /organism="Caenorhabditis elegans"
                        /mol_type="mRNA"
                        /strain="N2"
                        /db_xref="taxon:6239"
                        /clone="yk1029h03"
                        /sex="hermaphrodite"
                        /tissue_type="whole animal"
                        /dev_stage="L1"
                        /clone_lib="unpublished oligo-capped cDNA library, C.
                        elegans L1 stage"

BASE COUNT  233 a 150 c 144 g 232 t 2 others

ORIGIN
Query Match 39.5%; Score 591; DB 12; Length 761;
Best Local Similarity 96.6%; Pred. No. 1.4e-108;
Matches 622; Conservative 0; Mismatches 1; Indels 21; Gaps 1;

QY      854 ACTCAGACGACGTCACATTGATCATTTAGCAAAAGATCCATTGATCCCATGATTGGA 913
        |||||||
Db      761 ACTCAGACGACGTCACATTGATCATTTAGCAAAAGATCCATTGATCCCATGATTGGA 702
QY      914 GAAAGTTAAGTTGGACAAAGTTGAGCTCATGCACCCTGGCTCAGCATTTCAAT 973
        |||||||
Db      701 GAAAGTTAAGTTGGACAAAGTTGAGCTCATGCACCCTGGCTCAGCATTTCAAT 642
QY      974 ACCTCACGTCGCTTCGATTCGAAATTTGCAAAAGTGAAGATATCTTATGTGGAA 1033
        |||||||
Db      641 ACCTCACGTCGCTTCGATTCGAAATTTGCAAAAGTGAAGATATCTTATGTGGAA 582
QY      1034 TGATAGTCCAGATGACATTTGAAGACAGTTTCTTATCATATCAATTAATACATTATGT 1093
        |||||||
Db      581 TGATAGTCCAGATGACATTTGAAGACAGTTTCTTATCATATCAATTAATACATTATGT 522
QY      1094 TCCCACTTGTTATGGGAAAAGTATATTTGGAATGTTGCCGACAGATGAGTTGAAG 1153
        |||||||
Db      521 TCCCACTTGTTATGGGAAAAGTATATTTGGAATGTTGCCGACAGATGAGTTGAAG 462
QY      1154 GAACATTGAGATGATGATCTTGAGAGAAAGAAATCTGCAGAAACCTACCGCTTGACT 1213
        |||||||
Db      461 GAACATTGAGATGATGATCTTGAGAGAAAGAAATCTGCAGAAACCTACCGCTTGACT 402
QY      1214 TGTTCAGGCAATGCTTCCCAAGAGAGATTAGACAAATTTAAGGTAATTCATTTCCA 1273
        |||||||
Db      401 TGTTCAGGCAATGCTTCCCAAGAGAGATTAGACAAATTTTA----- 359

```

```

QY      1274 AACGGTAGACTACACCTTTGAAGCTGTCACATGTGTGAAAATCAGTTTATTTGTCAG 1333
        |||||||
Db      358 -----GGTAGACTACACCTTTGAAGCTGTCACATGTGTGAAAATCAGTTTATTTGTCAG 303
QY      1334 CTACAGTGAATCAGTTTCATGGAAGACGTGAATAATGTCATTTGCGACGCTGGGATGAG 1393
        |||||||
Db      302 CTACAGTGAATCAGTTTCATGGAAGACGTGAATAATGTCATTTGCGACGCTGGGATGAG 243
QY      1394 AATTGATGACGTGTGATGTGACTCCCATATATTTCTACCGATAGGATGTTGGAAG 1453
        |||||||
Db      242 AATTGATGACGTGTGATGTGACTCCCATATATTTCTACCGATAGGATGTTGGAAG 183
QY      1454 CGCAGATTATGTTCTACAACTCCGAAAAAGTACATTATGA 1497
        |||||||
Db      182 CGCAGATTATGTTCTACAACTCCGAAAAAGTACATTATGA 139

RESULT 4
LOCUS   CB404801 607 bp mRNA linear EST 15-MAY-2003
DEFINITION OSTR028D3_1 AD-wrmcDNA Caenorhabditis elegans cDNA, mRNA sequence.
ACCESSION CB404801
VERSION   CB404801.1 GI:30746528
KEYWORDS EST.
SOURCE   Caenorhabditis elegans
ORGANISM Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
          ; Rhabditidae; Peleoderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 607)
          Reboul,J., Vaglio,P., Rual,J.F., JAMESCH,P., Martinez,M., Armstrong
          C.M., Li,S., Jacotot,L., Berth,N., Janky,R., Moore,T., Hudson
          J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
          Endress,G.A., Jena,S., Chevret,E., Papasotiropoulos,V., Tolias,P.P.,
          Doucette-Stamm,L., Hill,D.E. and Vidal,M.
          C. elegans ORFome version 1.1: experimental verification of the
          genome annotation and resource for proteome-scale protein
          expression
JOURNAL Nat. Genet., (2003) In press
COMMENT  Contact: Vidal M
          Marc Vidal Laboratory
          Dana Farber Cancer Institute
          1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
          Tel: 617 632 5180
          Fax: 617 632 5739
          Email: Marc.Vidal@dfci.harvard.edu
          Sequence tag of Gateway entry clones. The primers used were
          designed on the predicted protein encoding ORF. C. elegans ORFome
          cloning project : Contact david.hill@dfci.harvard.edu or
          marc.vidal@dfci.harvard.edu
          POLY-A-No.
FEATURES
    source             1..607
                        /organism="Caenorhabditis elegans"
                        /mol_type="mRNA"
                        /strain="N2"
                        /db_xref="taxon:6239"
                        /sex="Hermaphrodite and male"
                        /tissue_type="whole animal"
                        /dev_stage="mixed stage"
                        /clone_lib="AD-wrmcDNA"
                        /clone="The AD-wrmcDNA library was generated with poly(A)+
                        RNA isolated from both hermaphrodite and male N2 worms of
                        all larval stages, embryos, adults and dauers and the
                        subsequent generation of cDNAs by poly(A) priming. The
                        cDNAs were cloned into pC86"

BASE COUNT  175 a 131 c 113 g 188 t

ORIGIN
Query Match 37.6%; Score 563.4; DB 14; Length 607;
Best Local Similarity 96.3%; Pred. No. 4.8e-103;
Matches 606; Conservative 0; Mismatches 1; Indels 22; Gaps 2;

```

QY 843 TCAGATATGACTCAGACGACGTCACATTTGATCAATTTAGCAAAAGATCCATTTGATCC 902
|||||
DB 607 TCCAAGATATGACTCAGACGACGTCACATTTGATCAATTTAGCAAAAGATCCATTTGATCC 548
QY 903 CATGATTTGGAGAAAAGTTAAGGTTGGACAAAAGTTAGCTCATGACCCCTTGCTCA 962
DB 547 CATGATTTGGAGAAAAGTTAAGGTTGGACAAAAGTTAGCTCATGACCCCTTGCTCA 488
QY 963 GCAATTTCAATTAACCTCCACGTCGCTTCGATTCCTCAATTTTGCAAAAGTGAAGATATCT 1022
DB 487 GCAATTTCAATTAACCTCCACGTCGCTTCGATTCCTCAATTTTGCAAAAGTGAAGATATCT 428
QY 1023 TATTTGGGAATGATGATGCTCAGATGACCTTGAAAGAGTTTCCATATTCATCAATTA 1082
DB 427 TATTTGGGAATGATGATGCTCAGATGACCTTGAAAGAGTTTCCATATTCATCAATTA 368
QY 1083 TACATTTATGTTCCAGATGTTGTTATGCGGAAAAGTATATTTGGAACTTTCCGCCGA 1142
DB 367 TACATTTATGTTCCAGATGTTGTTATGCGGAAAAGTATATTTGGAACTTTCCGCCGA 308
QY 1143 TGAGTTCAAGAACATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1202
DB 307 TGAGTTCAAGAACATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 248
QY 1203 ACCGCTGACTGTTTCAAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTAT 1262
DB 247 ACCGCTGACTGTTTCAAGCCAAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGTAT 194
QY 1263 TCTGATTTCCAAACGCGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1322
DB 193 TCTGATTTCCAAACGCGTACGATGATGATGATGATGATGATGATGATGATGATGATGAT 149
QY 1323 TATTTGCCAGCTACAGTGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 1382
DB 148 TATTTGCCAGCTACAGTGAATATGATGATGATGATGATGATGATGATGATGATGATGAT 89
QY 1383 CTGGGATGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1442
DB 88 CTGGGATGAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 29
QY 1443 ATGGTGTGAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1471
DB 28 ATGGTGTGAGGCGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 5
Bj101701 581 bp mRNA linear EST 18-JAN-2002
LOCUS Bj101701 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1029h03 5', mRNA sequence.
ACCESSION Bj101701
VERSION Bj101701.1 GI:18244371
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
AUTHORS 1 (bases 1 to 581)
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..581
/organism="Caenorhabditis elegans"

/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1029h03"
/sex="hermaphrodite"
/russue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 189 a 103 c 123 g 166 t
ORIGIN
Query Match 37.3%, Score 559, DB 12, Length 581;
Best Local Similarity 100.0%; Pred. No. 3, 6e-102;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTGAAATTTGAAAATTTGAGAGCTTAACAAAATGAGAGAAATCTGATTAAG 60
DB 23 ATGCTGAAATTTGAAAATTTGAGAGCTTAACAAAATGAGAGAAATCTGATTAAG 82
QY 61 ACCTACTTTGGGAAATCTTATTTACATCAGTTGAGAAAAGAAAACCTTTTCATTCCA 120
DB 83 ACCTACTTTGGGAAATCTTATTTACATCAGTTGAGAAAAGAAAACCTTTTCATTCCA 142
QY 121 GTTGAGATTCATTCGTAACCTTACAGTTATTTTAAAGAAATGCGTGAAGAGAGTT 180
DB 143 GTTGAGATTCATTCGTAACCTTACAGTTATTTTAAAGAAATGCGTGAAGAGAGTT 202
QY 181 ATCTTGAAGACAGTGCATGATTAATGACAAGAACTGATGATGATGATGATGATGATGATGAT 240
DB 203 ATCTTGAAGACAGTGCATGATTAATGACAAGAACTGATGATGATGATGATGATGATGATGAT 262
QY 241 TTTGACGAAATTTGAAAATTTGGGATACAGAGTTTGGCTCAGTTTATGAGCTGAC 300
DB 263 TTTGACGAAATTTGAAAATTTGGGATACAGAGTTTGGCTCAGTTTATGAGCTGAC 322
QY 301 ACGAAATTTTGGCTCAATTTTATCGGACGATATGTTGGTTTGGCAAAAGCCGCAATG 360
DB 323 ACGAAATTTTGGCTCAATTTTATCGGACGATATGTTGGTTTGGCAAAAGCCGCAATG 382
QY 361 AGTATCCCATTAATGATTAATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 420
DB 383 AGTATCCCATTAATGATTAATGATATGATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 421 CAAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 480
DB 443 CAAAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAATGATTAAT 502
QY 481 TCGCTGCTCCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
DB 503 TCGCTGCTCCAAAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 562
QY 541 GGACAACTCTTGAACATAT 559
DB 563 GGACAACTCTTGAACATAT 581

RESULT 6
Bj124772 557 bp mRNA linear EST 30-MAY-2003
LOCUS Bj124772 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1315b11 5', mRNA sequence.
ACCESSION Bj124772
VERSION Bj124772.2 GI:31246008
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
AUTHORS 1 (bases 1 to 557)
TITLE A complementary view of the C. elegans genome

JOURNAL Unpublished
 COMMENT On Jan 23, 2002 this sequence version replaced gi:18284913.
 Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers
 1.557
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK1315b11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 180 a 99 c 118 g 158 t 2 others
 ORIGIN
 Query Match 34.8%; Score 521; DB 12; Length 557;
 Best Local Similarity 99.8%; Pred. No. 1.6e-94;
 Matches 532; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 3 GTCGTGAATTCGAAATTTGTCAGAGCTTAACAAAAATCGGACGAAACTCGATTAAGC 62
 DB 25 GTCGTGAATTCGAAATTTGTCAGAGCTTAACAAAAATCGGACGAAACTCGATTAAGC 84
 QY 63 CTACTGTGGGAATCCATTTTACATCAGTTCGAGAAAG-AAAACCTTCTTCATTCGAG 121
 DB 85 CTACTGTGGGAATCCATTTTACATCAGTTCGAGAAAGGAAAACTTCTTCATTCGAG 144
 QY 122 TTGAAGCATTCATCGTAACCTTACGTTAATTTTAAAGATCCGTGAAGGAAGAGTTA 181
 DB 145 TTGAAGCATTCATCGTAACCTTACGTTAATTTTAAAGATCCGTGAAGGAAGAGTTA 204
 QY 182 TCTTGAACAGTGGTCCATGATTAAGCAAGAACCTGCGATTCGATTCAGTCAAGTGGT 241
 DB 205 TCTTGAACAGTGGTCCATGATTAAGCAAGAACCTGCGATTCGATTCAGTCAAGTGGT 264
 QY 242 TTGACAGATTTGAAAGTTTGCGATACAGAGTTTGCGTCAAGTTTATCGAGCTGACA 301
 DB 265 TTGACAGATTTGAAAGTTTGCGATACAGAGTTTGCGTCAAGTTTATCGAGCTGACA 324
 QY 302 CGAAATTTTGGCTCAATATTTTATCGAGATATGTTTGGTTGGCAACGCCGCAATGA 361
 DB 325 CGAAATTTTGGCTCAATATTTTATCGAGATATGTTTGGTTGGCAACGCCGCAATGA 384
 QY 362 GTGATCCCAATATGATAAATTTGTATATGCTCCGCCGCTTGCAATCAACGAAGATACC 421
 DB 385 GTGATCCCAATATGATAAATTTGTATATGCTCCGCCGCTTGCAATCAACGAAGATACC 444
 QY 422 AAAATGATATGTAATTTGTAATTAATTTGCAATTCGCGAAATCGTGGGCAAACTT 481
 DB 445 AAAATGATATGTAATTTGTAATTAATTTGCAATTCGCGAAATCGTGGGCAAACTT 504
 QY 482 CGCTGTCTCCAAATTCGATGAAGGAGGCTCTCTCTAAGCAAGCATCGTTTC 534
 DB 505 CGCTGTCTCCAAATTCGATGAAGGAGGCTCTCTCTAAGCAAGCATCGTTTC 557

RESULT 7
 AU216908/c 677 bp mRNA linear EST 17-JUL-2001
 LOCUS AU216908 unpublished oligo-capped cDNA library, stage I4
 DEFINITION Caenorhabditis elegans cDNA clone yk843e05 3', mRNA sequence.
 ACCESSION AU216908
 VERSION AU216908.1 GI:14855065
 KEYWORDS EST.

SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae
 1 (bases 1 to 677)
 REFERENCE Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.
 A complementary view of the C.elegans genome
 TITLE Unpublished
 JOURNAL Contact: Yoji Kohara
 COMMENT Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.
 Location/Qualifiers
 1.677
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK843e05"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L4"
 /clone_lib="unpublished oligo-capped cDNA library, stage
 L4"
 BASE COUNT 210 a 136 c 130 g 201 t
 ORIGIN
 Query Match 34.7%; Score 519; DB 9; Length 677;
 Best Local Similarity 96.3%; Pred. No. 4e-94;
 Matches 550; Conservative 0; Mismatches 0; Indels 21; Gaps 1;

QY 927 TGGCAAAAGTTTGACCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGC 986
 DB 677 TGGCAAAAGTTTGACCTCATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGC 618
 QY 987 TTGCATTCGAAATTTTGAAGAACTGAAGATATCTATTGGGGAATGATGGTCCACA 1046
 DB 617 TTGCATTCGAAATTTTGAAGAACTGAAGATATCTATTGGGGAATGATGGTCCACA 558
 QY 1047 TGCACCTTGAAGACAGTTTCTATTTCATATCAATTAATACATTATGTTCCAGTTGTTA 1106
 DB 557 TGCACCTTGAAGACAGTTTCTATTTCATATCAATTAATACATTATGTTCCAGTTGTTA 498
 QY 1107 TGGGAAAGATATTTGGAACCTGTTCCGCCAGATGAGTCAAGAACATTCAGATG 1166
 DB 497 TGGGAAAGATATTTGGAACCTGTTCCGCCAGATGAGTCAAGAACATTCAGATG 438
 QY 1167 GGATGATACTCTGGCAAAAGATCTGCAGAAACCCATACCGCTTGACTTGTCAAGCCAT 1226
 DB 437 GGATGATACTCTGGCAAAAGATCTGCAGAAACCCATACCGCTTGACTTGTCAAGCCAT 378
 QY 1227 GCCTTCCCAAGAGAGATTAGACAAATTTAAGTAATTCGATTTCGAAACGGGTAGACT 1286
 DB 377 GCCTTCCCAAGAGAGATTAGACAAATTTA-----GGTAGACT 339
 QY 1287 ACGCTTGAAGCTGTGACATGTGGAATAATCAGTTTATTTGTCCAGCTACAGTGAATC 1346
 DB 338 ACGCTTGAAGCTGTGACATGTGGAATAATCAGTTTATTTGTCCAGCTACAGTGAATC 279
 QY 1347 AGTTCAATGAACATCATTAATAGTCAATTTGCAAGCGCTGGATGGAAGATTTGATCAACT 1406
 DB 278 AGTTCAATGAACATCATTAATAGTCAATTTGCAAGCGCTGGATGGAAGATTTGATCAACT 219
 QY 1407 GTATGATGTGACCTCCATGATATTTCTACAGATAGATGTTGGAAGCCACAGATTATGT 1466
 DB 218 GTATGATGTGACCTCCATGATATTTCTACAGATAGATGTTGGAAGCCACAGATTATGT 159
 QY 1467 TCTACACCTCCGAAAAAGTACACTATTGA 1497
 |||||||

Db 158 TCACACCTCCGAAAAAGTACACTATGCA 128

RESULT 8
Bj153087/c
LOCUS
DEFINITION Bj153087 686 bp mRNA linear EST 24-JAN-2002
Bj153087 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone YK1315B11 3', mRNA sequence.
ACCESSION
VERSION Bj153087.1 GI:18321072
KEYWORDS
SOURCE EST.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
; Rhabdilitidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
1 (bases 1 to 686)
A complementary view of the C.elegans genome
TITLE
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..686
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1315B11"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 214 a 133 c 131 g 204 t 4 others
ORIGIN

Query Match 34.1%; Score 510; DB 12; Length 686;
Best Local Similarity 96.1%; Pred. No. 2.6e-92;
Matches 541; Conservative 0; Mismatches 1; Indels 21; Gaps 1;

QY 935 AGTTGAGCTCATGACCCCTTGGCTCAGCAATTCATACCTCAGCTCGCTTCGATTC 994
|||||
Db 686 AGTTGAGCTCATGACCCCTTGGCTCAGCAATTCATACCTCAGCTCGCTTCGATTC 627
|||||

QY 995 TCAATATTTGCAAACTGAGAGATATCTTATTGGGGAATGAGATGCTCCAGATGCACCTTG 1054
|||||
Db 626 TCAATATTTGCAAACTGAGAGATATCTTATTGGGGAATGAGATGCTCCAGATGCACCTTG 567
|||||

QY 1055 AAGCAGTTTCCTATTCATATCAATATATACATTATGTCCTCCAGTTGGTATGGGAAA 1114
|||||
Db 566 AAGCAGTTTCCTATTCATATCAATATATACATTATGTCCTCCAGTTGGTATGGGAAA 507
|||||

QY 1115 AGTATATTTGGAAGCTTGTCCGCCAGATGATTCAAAGAAACATTCAGATGGATGAT 1174
|||||
Db 506 AGTATATTTGGAAGCTTGTCCGCCAGATGATTCAAAGAAACATTCAGATGGATGAT 447
|||||

QY 1175 ACTTGAGAAAGATCTGAGAAACCTTACCGCTTGACTGTTGCAAGCAATGCTTCCC 1234
|||||
Db 446 ACTTGAGAAAGATCTGAGAAACCTTACCGCTTGACTGTTGCAAGCAATGCTTCCC 387
|||||

QY 1235 AAGGAGATTGACAAATTTAAGTAATTGATTTCCAAACGGGTAGAGACTAGCCCTTG 1294
|||||
Db 386 AAGGAGATTGACAAATTTA-----GCTAGACTAGCCCTTG 348
|||||

QY 1295 AAGCTGTCAGATGTGTAAATCAGTTATTTGTCCAGCTACAGATGAATCAGTTCAATG 1354
|||||
Db 347 AAGCTGTCAGATGTGTAAATCAGTTATTTGTCCAGCTACAGATGAATCAGTTCAATG 288
|||||

QY 1355 GAAGACTGATTAATGTCATTTTGGACGGCTGGGATGAGAATTTGATGAACGTATGATG 1414
|||||
Db 287 GAAGACTGATTAATGTCATTTTGGACGGCTGGGATGAGAATTTGATGAACGTATGATG 228
|||||

QY 1415 TGGACCTCCATGATATCTACCGATAGAGATGTTGTAAGCGCACAGTTATGTTCTACAC 1474
|||||
Db 227 TGGACCTCCATGATATCTACCGATAGAGATGTTGTAAGCGCACAGTTATGTTCTACAC 168
|||||

QY 1475 CTCGAAAAAGTACACTATGCA 1497
|||||
Db 167 CTCGAAAAAGTACACTATGCA 145
|||||

RESULT 9
Bj115564
LOCUS
DEFINITION Bj115564 519 bp mRNA linear EST 23-JAN-2002
Bj115564 unpublished oligo-capped cDNA library, C. elegans L1 stage
Caenorhabditis elegans cDNA clone YK1192H09 5', mRNA sequence.
ACCESSION
VERSION Bj115564.1 GI:18275658
KEYWORDS
SOURCE EST.
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
; Rhabdilitidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS Kohara,Y., Shih-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
1 (bases 1 to 519)
A complementary view of the C.elegans genome
TITLE
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
Location/Qualifiers
1..519
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1192H09"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 170 a 91 c 109 g 149 t
ORIGIN

Query Match 33.6%; Score 502.4; DB 12; Length 519;
Best Local Similarity 99.8%; Pred. No. 8.5e-91;
Matches 503; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGCTGATTTTGAATAATTTGTAGAGCTTAACAAAAATGAGCAGAAAACTGGATTAAG 60
|||||
Db 16 ATGCTGATTTTGAATAATTTGTAGAGCTTAACAAAAATGAGCAGAAAACTGGATTAAG 75
|||||

QY 61 ACCTACTGTGGGAATCCTATTATACATCAGTTTCAGAAAGAAAAAATCTTCTTCAATCCA 120
|||||
Db 76 ACCTACTGTGGGAATCCTATTATACATCAGTTTCAGAAAGAAAAAATCTTCTTCAATCCA 135
|||||

QY 121 GTTGAAGCATTCATTCGTACCTTACAGTTAATTTTAAAGATCGTGAAGAAAGAGACTT 180
|||||
Db 136 GTTGAAGCATTCATTCGTACCTTACAGTTAATTTTAAAGATCGTGAAGAAAGAGACTT 195
|||||

QY 181 ATCTTGAAGAAAGTGTGATGATTAATGACAAAGAACTCGATTTGATTCATCACTCAGATG 240
|||||
Db 196 ATCTTGAAGAAAGTGTGATGATTAATGACAAAGAACTCGATTTGATTCATCACTCAGATG 255
|||||

```
QY 241 TTTCACGAATGAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGAC 300
    |||||||
DB 256 TTTCACGAATGAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGGAGCTGAC 315
QY 301 ACCAAATTTTGGCTCAATATTTATCGGACGATATGTTGGTTGGCAAAAGCCGCAATG 360
    |||||||
DB 316 ACCAAATTTTGGCTCAATATTTATCGGACGATATGTTGGTTGGCAAAAGCCGCAATG 375
QY 361 AGGATCCCAATATGATTAATTTATGATGCTCCGCCGCTTCATTCACGAAGAATAC 420
    |||||||
DB 376 AGGATCCCAATATGATTAATTTATGATGCTCCGCCGCTTCATTCACGAAGAATAC 435
QY 421 CAAATGATATGTTAAATATGTAATTAATTCATTCATTCATTCATTCATTCATTCATTC 480
    |||||||
DB 436 CAAATGATATGTTAAATATGTAATTAATTCATTCATTCATTCATTCATTCATTCATTC 495
QY 481 TCGCTGTCTCCAAATTCGATGAA 504
    |||||||
DB 496 TCGCTGTCTCCAAATTCGATGAA 519

RESULT 10
AV197399 380 bp mRNA linear EST 26-JUL-1999
LOCUS AV197399 yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
DEFINITION embryo Caenorhabditis elegans cDNA clone yk65h11 5', mRNA
sequence.
ACCESSION AV197399.1 GI:5581170
VERSION AV197399.1
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 380)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,
Nishigaki,A., Motohashi,T., Zeng,Q., Watanabe,H., Sugimoto,A., Sano
,M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and
Nomoto,H.
TITLE Expressed genes in C.elegans
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..380
Location/Qualifiers
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk65h11"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone_lib="yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 125 a 63 c 83 g 108 t 1 others
ORIGIN
Query Match 24.4%; Score 366; DB 9; Length 380;
Best Local Similarity 99.7%; Pred. No.2,2e-63;
Matches 377; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
```

```
QY 135 TCGTAACTTACAGTTAATTTTAACGAATGCGTGAAGAAAGAGTTATCTTGAACAGT 194
    |||||||
DB 123 TCGTAACTTACAGTTAATTTTAACGAATGCGTGAAGAAAGAGTTATCTTGAACAGT 182
QY 195 GGTCCATGATTAAGACAGAAGCTCGATTCGATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 254
    |||||||
DB 183 GGTCCATGATTAAGACAGAAGCTCGATTCGATTCAGTTCAGTTCAGTTCAGTTCAGTTCAG 242
QY 255 AAAAGTTTGGCGATACAGAGTTTGGCTCAGTTTATCGGAGCTGACACGAAATTTGGCT 314
    |||||||
DB 243 AAAAGTTTGGCGATACAGAGTTTGGCTCAGTTTATCGGAGCTGACACGAAATTTGGCT 302
QY 315 CAATATTTATCGGACGATATGTTGGTTGGCAAAAGCCGCAATGATGATCCATAT 374
    |||||||
DB 303 CAATATTTATCGGACGATATGTTGGTTGGCAAAAGCCGCAATGATGATCCATAT 362
QY 375 GGATAAATTTGATATGC 392
    |||||||
DB 363 GGATAAATTTGATATGC 380

RESULT 11
C42305 360 bp mRNA linear EST 18-OCT-1999
LOCUS C42305 yuji Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
DEFINITION Caenorhabditis elegans cDNA clone yk292d9 5', mRNA sequence.
ACCESSION C42305.1 GI:2378542
VERSION C42305.1
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 360)
AUTHORS Kohara,Y., Motohashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
,M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1..360
Location/Qualifiers
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk292d9"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone_lib="yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 119 a 62 c 78 g 101 t
ORIGIN
Query Match 24.0%; Score 360; DB 14; Length 360;
Best Local Similarity 100.0%; Pred. No.3,4e-62;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 24 CAGAGCTAACAAAAATCGACAGAAAACCTCGATAAGACCTACTTGTGGGAATCCTATT 83
    |||||||
DB 1 CAGAGCTAACAAAAATCGACAGAAAACCTCGATAAGACCTACTTGTGGGAATCCTATT 60
QY 84 ACATCATTCGAGAAAGAAAACCTCTTCATTCGATTCGATTCGATTCGATTCGATTCGATTC 143
    |||||||
DB 61 ACATCATTCGAGAAAGAAAACCTCTTCATTCGATTCGATTCGATTCGATTCGATTCGATTC 120
QY 144 TACAGTTAATTTTACGAATGCGTGAAGAAAGAGTTATCTTGAAGAAAGTGTGCATGA 203
```

```
|||||
Db 121 TACAGTATATTTTAAACGATGCGTGAAGAGAGATTATCTTCGAAACAGGGTCATGA 180
Qy 204 TTATGACAAAGAACTGCGATTGCAATGATGTTTGCAGCAATTTGAAAAGTTTG 263
Db 181 TTATGACAAAGAACTGCGATTGCAATGATGTTTGCAGCAATTTGAAAAGTTTG 240
Qy 264 CGGATPACAGAGCTTCTGCTCAGTTTATCGAGCTGACAGAAATTTTGGCTCAATATTTT 323
Db 241 CGGATPACAGAGCTTCTGCTCAGTTTATCGAGCTGACAGAAATTTTGGCTCAATATTTT 300
Qy 324 ATCGAGCATATGTTTGGTTGGCAACGCCGCAATGATGATCCCATATGATGATTAAT 383
Db 301 ATCGAGCATATGTTTGGTTGGCAACGCCGCAATGATGATCCCATATGATGATTAAT 360

RESULT 12
LOCUS AU111588 340 bp mRNA linear EST 19-OCT-2000
DEFINITION AU111588 unpublished oligo-capped cDNA library Caenorhabditis
ACCESSION AU111588
VERSION AU111588
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
AUTHORS ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 340)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
National Institute of Genetics
Genome Biology Lab.
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1. .340
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 111 a 58 c 72 g 98 t 1 others
ORIGIN
Query Match 20.6%; Score 308; DB 9; Length 340;
Best Local Similarity 100.0%; Pred. No. 9.7e-52;
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 GTCGAAATTTTGAATAATGTCAGAGCTAACAAAAATCGACAGAAAACCTGATTAAC 62
Db 33 GTCGAAATTTTGAATAATGTCAGAGCTAACAAAAATCGACAGAAAACCTGATTAAC 92
Qy 63 CTACTTGGGAATCTATATTACATCACTGCGAGAGAAAACCTTCTTCAATCCAGT 122
Db 93 CTACTTGGGAATCTATATTACATCACTGCGAGAGAAAACCTTCTTCAATCCAGT 152
Qy 123 TGAAGCATTCATCTTAACCTTACAGTTAATTTTAAACAATGCCGGAAGAGAGTTAT 182
Db 153 TGAAGCATTCATCTTAACCTTACAGTTAATTTTAAACAATGCCGGAAGAGAGTTAT 212
Qy 183 CTTCGAACAGTGTGTCATGATTAAGACAGAACTGCAATTCGATTCAGTCAGATGTT 242
Db 213 CTTCGAACAGTGTGTCATGATTAAGACAGAACTGCAATTCGATTCAGTCAGATGTT 272
```

```
Qy 243 TGCACGAATTTGAAAAGTTTCCGATACAGAGTTCTGCTCAGTTTATCGAGCTGACAC 302
Db 273 TGCACGAATTTGAAAAGTTTCCGATACAGAGTTCTGCTCAGTTTATCGAGCTGACAC 332
Qy 303 GAAATTTT 310
Db 333 GAAATTTT 340

RESULT 13
LOCUS AU116296/c 300 bp mRNA linear EST 19-OCT-2000
DEFINITION AU116296 unpublished oligo-capped cDNA library Caenorhabditis
ACCESSION AU116296
VERSION AU116296
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabditoidea
AUTHORS ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
National Institute of Genetics
Genome Biology Lab.
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
FEATURES
source
1. .300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 95 a 63 c 54 g 88 t
ORIGIN
Query Match 14.7%; Score 220; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 4.6e-34;
Matches 220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1278 GGTAGGACTACGCCCTTGAAGCTGTGCATGTGTGAATAATCATTTATTTTCCAGCTAC 1337
Db 300 GGTAGGACTACGCCCTTGAAGCTGTGCATGTGTGAATAATCATTTATTTTCCAGCTAC 241
Qy 1338 AGTGAATCAAGTTCAATGGAACACTGATTAATGTCAATTTCCAGCGCTGGATGAAGATT 1397
Db 240 AGTGAATCAAGTTCAATGGAACACTGATTAATGTCAATTTCCAGCGCTGGATGAAGATT 181
Qy 1398 TGAGGAAGCTGATGATGTGGAGCTCCCATGATTTCTTACGATAGGATGGTGAAGCGCA 1457
Db 180 TGAGGAAGCTGATGATGTGGAGCTCCCATGATTTCTTACGATAGGATGGTGAAGCGCA 121
Qy 1458 CAGTTATGTTCTACAACTCCGAAAAGATACACTATTGA 1497
Db 120 CAGTTATGTTCTACAACTCCGAAAAGATACACTATTGA 81

RESULT 14
LOCUS AU115428 300 bp mRNA linear EST 19-OCT-2000
DEFINITION AU115428 unpublished oligo-capped cDNA library Caenorhabditis
```

Accession A0115428
Version A0115428.1 GI:10928995
KeyWords EST.
Source Caenorhabditis elegans
Organism Caenorhabditis elegans
Reference Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabdilitida; Rhabditoidea
Authors ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 300)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 94 a 60 c 54 g 89 t 3 others
ORIGIN

Query Match 14.4%; Score 215; DB 9; Length 300;
Best Local Similarity 100.0%; Pred. No. 4,7e-33;
Matches 215; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1283 GACTACGCGCTGAGAGCTGCTACATGCTGAAATTCATTTGTCACGCTACAGTCA 1342
|||||
DB 300 GACTACGCGCTGAGAGCTGCTACATGCTGAAATTCATTTGTCACGCTACAGTCA 241
|||||
OY 1343 AATCAGTTTCATGAGAGCTGATTAATGTCATTTTCGAGCGCTGGATGAGAAATTTGATG 1402
|||||
DB 240 AATCAGTTTCATGAGAGCTGATTAATGTCATTTTCGAGCGCTGGATGAGAAATTTGATG 181
|||||
OY 1403 AACTGTATGATGTGACTCCCATGATTTTACCGATAGATGCTGTGAGCGACAGTT 1462
|||||
DB 180 AACTGTATGATGTGACTCCCATGATTTTACCGATAGATGCTGTGAGCGACAGTT 121
|||||
OY 1463 ATGTTCTACACCTCCGAAAGTACAACTATTGA 1497
|||||
DB 120 ATGTTCTACACCTCCGAAAGTACAACTATTGA 86
|||||

RESULT 15
A0112450 337 bp mRNA linear EST 30-MAY-2003
LOCUS A0112450
DEFINITION A0112450 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk743a4 5', mRNA sequence.
ACCESSION A0112450
VERSION A0112450.2 GI:31237498
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabdilitida; Rhabditoidea
Authors ; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 337)
Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT On Oct 19, 2000 this sequence version replaced gi:10926017.

Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1..337
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN

Query Match 12.4%; Score 186.2; DB 9; Length 337;
Best Local Similarity 97.9%; Pred. No. 2.9e-27;
Matches 188; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 ATGCTGCAATTTTCGAAATTTGTCAGAGCTAACAAGAAATCGACAGAAATCGATPAG 60
|||||
DB 23 ATGCTGCAATTTTCGAAATTTGTCAGAGCTAACAAGAAATCGACAGAAATCGATPAG 82
|||||
OY 61 ACCACTGTGCGAATCCTATTTCATCAGTTCGAGAAAGAAACCTTCTTCATTCCA 120
|||||
DB 83 ACCACTGTGCGAATCCTATTTCATCAGTTCGAGAAAGAAACCTTCTTCATTCCA 142
|||||
OY 121 GTTGAAGATTCAATCGTAACCTTACAGTTAATTTTAAACGAATGCGTGAAGAGAGTT 180
|||||
DB 143 GTTGAAGATTCAATCGTAACCTTACAGTTAATTTTAAACGAATGCGTGAAGAGAGTT 202
|||||
OY 181 ATCTTCGAAACA 192
|||||
DB 203 ATCTTCGAAACA 214
|||||

Search completed: September 2, 2003, 17:41:43
Job time : 3449 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 17:41:55 ; Search time 3440 Seconds
(without alignments)
10576.686 Million cell updates/sec

Title: US-09-872-523-6
Perfect score: 1497
Sequence: 1 atgtctgaattcttgaataa.....cgaaaagtacacatttga 1497

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 22781392 seqs, 12152238056 residues

Word size : 0

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

EST:*
1: em_estha:*
2: em_esthum:*
3: em_estin:*
4: em_estimu:*
5: em_estrov:*
6: em_estropl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	581	38.8	653	9 AU204939	AU204939 AU204939
2	559	37.3	581	12 BJ101701	BJ101701 BJ101701
3	453	30.3	519	12 BJ115564	BJ115564 BJ115564
4	433	28.9	557	12 BJ124772	BJ124772 BJ124772

C 5	414	27.7	768	12	BJ143186	BJ143186 BJ143186
C 6	385	25.7	607	14	CB404801	CB404801 OSTR028D3
C 7	360	24.0	360	14	CA4305	CA4305 CA4305 Yuj1
C 8	356	23.8	761	12	BJ129361	BJ129361 BJ129361
C 9	334	22.3	677	9	AU216908	AU216908 AU216908
C 10	322	21.5	380	9	AV197399	AV197399 AV197399
C 11	308	20.6	340	9	AU111588	AU111588 AU111588
C 12	306	20.4	686	12	BJ153087	BJ153087 BJ153087
C 13	220	14.7	300	9	AU116296	AU116296 AU116296
C 14	215	14.4	300	9	AU115428	AU115428 AU115428
C 15	186	12.4	337	9	AU112450	AU112450 AU112450
C 16	66	4.4	505	14	CB385690	CB385690 OSTR028D3
C 17	23	1.5	1291	10	BC503414	BC503414 602550841
C 18	22	1.5	265	12	BM632811	BM632811 170006875
C 19	22	1.5	304	14	CB698306	CB698306 AMGNNUC:N
C 20	22	1.5	426	14	CB695113	CB695113 AMGNNUC:N
C 21	22	1.5	507	28	AZ907625	AZ907625 RPT1-24-1
C 22	22	1.5	555	14	CB545230	CB545230 AMGNNUC:N
C 23	21	1.4	335	14	H34557	H34557 EST111592 R
C 24	21	1.4	399	28	BH787508	BH787508 FZMB020F0
C 25	21	1.4	453	14	CB395614	CB395614 OSTR159D8
C 26	21	1.4	467	10	BE416260	BE416260 MUG006.C1
C 27	21	1.4	467	10	BE416260	BE416260 MUG006.C1
C 28	21	1.4	493	29	CC063342	CC063342 fyma00140
C 29	21	1.4	635	12	BJ290111	BJ290111 BJ290111
C 30	21	1.4	638	29	BZ724245	BZ724245 PUDAX807D
C 31	21	1.4	758	28	A2124760	A2124760 OSJNB007
C 32	21	1.4	776	14	CB863916	CB863916 AGCNCOURT
C 33	21	1.4	839	13	B0358433	B0358433 603478104
C 34	21	1.4	872	9	AU121368	AU121368 AU121368
C 35	21	1.4	1032	29	BZ824126	BZ824126 PUGB047D
C 36	20	1.3	283	10	BF525236	BF525236 UI-R-AB1-
C 37	20	1.3	333	28	BH555727	BH555727 BOHNM36TF
C 38	20	1.3	360	9	AV194140	AV194140 AV194140
C 39	20	1.3	360	14	CA45578	CA45578 CA45578 Yuj1
C 40	20	1.3	360	14	D74616	D74616 CELK083CXF
C 41	20	1.3	385	9	AM640232	AM640232 b192n09.w
C 42	20	1.3	404	29	BZ491812	BZ491812 BONCO12FR
C 43	20	1.3	405	14	T94263	T94263 ye31f02.r1
C 44	20	1.3	420	14	CB798312	CB798312 AMGNNUC:N
C 45	20	1.3	430	13	B0804167	B0804167 SUPBHC09

ALIGNMENTS

RESULT 1
AU204939
LOCUS
DEFINITION
AU204939 unpublished oligo-capped cDNA library, stage LA
Caenorhabditis elegans cDNA clone yk843e05 5', mRNA sequence.
ACCESSION
AU204939
VERSION
AU204939.1 GI:14836792
KEYWORDS
EST.
SOURCE
Caenorhabditis elegans
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae
; Rhabditidae; Pelodderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 653)
Kohara,Y., Shih-L.T., Thierry-Mieg,J., Suzuki,Y.
and Sugano,S.
A complementary view of the C.elegans genome
TITLE
JOURNAL
COMMENT
Unpublished
Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 653
/organism="Caenorhabditis elegans"

```
/mol_type="mRNA"
/db_xref="taxon:6239"
/clone="YK843e05"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cdna library, stage
L4"

BASE COUNT      216 a      116 c      136 g      185 t
ORIGIN
Query Match      38.8%; Score 581; DB 9; Length 653;
Best Local Similarity 99.8%; Pred. No. 9.7e-303;
Matches 631; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 60
    |||||
DB 22 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 81
    |||||
OY 61 ACCTACTTGTGGGAATCCTATTATACATCAGTTGCGAAGAAAACCTTCTTCA 120
    |||||
DB 82 ACCTACTTGTGGGAATCCTATTATACATCAGTTGCGAAGAAAACCTTCTTCA 141
    |||||
OY 121 GTTGAACGATTCATCTGAACCTTACGATTAATTTTAACGAATCCGTGAAGAGAGT 180
    |||||
DB 142 GTTGAACGATTCATCTGAACCTTACGATTAATTTTAACGAATCCGTGAAGAGAGT 201
    |||||
OY 181 ATCTTGAACAGAGTGTCATGATTATGACAAAGTCGATTCGATTCAAGTCAGATG 240
    |||||
DB 202 ATCTTGAACAGAGTGTCATGATTATGACAAAGTCGATTCGATTCAAGTCAGATG 261
    |||||
OY 241 TTTCACAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 300
    |||||
DB 262 TTTCACAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 321
    |||||
OY 301 ACGAAATTTTGGCCCAATTTTATCGAGCATATGTTGGTTGGCAAAAGCCGCAATG 360
    |||||
DB 322 ACGAAATTTTGGCCCAATTTTATCGAGCATATGTTGGTTGGCAAAAGCCGCAATG 381
    |||||
OY 361 AGTATCCCAATATGATTAATTTATATGCTCCGCCGCTTGCATCAACGAAGATAC 420
    |||||
DB 382 AGTATCCCAATATGATTAATTTATATGCTCCGCCGCTTGCATCAACGAAGATAC 441
    |||||
OY 421 CAAAATGATATGTAATTTATGTAATATTCATGATGCGCAATTCGCGCAACT 480
    |||||
DB 442 CAAAATGATATGTAATTTATGTAATATTCATGATGCGCAATTCGCGCAACT 501
    |||||
OY 481 TCGCTGCTCCAAAATTCGATGAAGGAGCTCTCCTAAGCAAGCATCGTTCAAGTT 540
    |||||
DB 502 TCGCTGCTCCAAAATTCGATGAAGGAGCTCTCCTAAGCAAGCATCGTTCAAGTT 561
    |||||
OY 541 GGACAAGCTCTTGAACATTTTAAATTTATTCCAATTTCTACTGAATACGGGTACGGCAAT 600
    |||||
DB 562 GGACAAGCTCTTGAACATTTTAAATTTATTCCTCAATTTCTACTGAATACGGGTACGGCAAT 621
    |||||
OY 601 CAAGAATATATGTAAGCAAGCAATGATATC 632
    |||||
DB 622 CAAGAATATATGTAAGCAAGCAATGATATC 653
    |||||

RESULT 2
BUI01701      581 bp      mRNA      linear      EST 18-JAN-2002
LOCUS      BUI01701 unpublished oligo-capped cdna library, C. elegans L1 stage
DEFINITION      Caenorhabditis elegans cdna clone yk1029h03 5', mRNA sequence.
ACCESSION      BUI01701
VERSION      BUI01701.1 GI:18244371
KEYWORDS      EST.
SOURCE      Caenorhabditis elegans
ORGANISM      Caenorhabditis elegans
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
            ; Rhabditidae; Peloderinae; Caenorhabditis.
```

```
REFERENCE      1 (bases 1 to 581)
AUTHORS      Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
TITLE      A complementary view of the C.elegans genome
JOURNAL      Unpublished
COMMENT      Contact: Tadasi Shin-I
            Center For Genetic Resource Information
            National Institute of Genetics
            111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshin@genetics.nig.ac.jp.
            Location/Qualifiers
FEATURES
    source
    1..581
    /organism="Caenorhabditis elegans"
    /mol_type="mRNA"
    /strain="N2"
    /db_xref="taxon:6239"
    /clone="YK1029h03"
    /sex="hermaphrodite"
    /tissue_type="whole animal"
    /dev_stage="L1"
    /clone_lib="unpublished oligo-capped cdna library, C.
    elegans L1 stage"

BASE COUNT      189 a      103 c      123 g      166 t
ORIGIN
Query Match      37.3%; Score 559; DB 12; Length 581;
Best Local Similarity 100.0%; Pred. No. 7.7e-291;
Matches 559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 60
    |||||
DB 23 ATGCTGAATTTTCGAAATTTGTCAGAGCTACACAAAATCGACAGAAAACCTCGATAG 82
    |||||
OY 61 ACCTACTTGTGGGAATCCTATTATACATCAGTTGCGAAGAAAACCTTCTTCA 120
    |||||
DB 83 ACCTACTTGTGGGAATCCTATTATACATCAGTTGCGAAGAAAACCTTCTTCA 142
    |||||
OY 121 GTTGAACGATTCATCTGAACCTTACGATTAATTTTAACGAATCCGTGAAGAGAGT 180
    |||||
DB 143 GTTGAACGATTCATCTGAACCTTACGATTAATTTTAACGAATCCGTGAAGAGAGT 202
    |||||
OY 181 ATCTTGAACAGAGTGTCATGATTATGACAAAGTCGATTCGATTCAAGTCAGATG 240
    |||||
DB 203 ATCTTGAACAGAGTGTCATGATTATGACAAAGTCGATTCGATTCAAGTCAGATG 262
    |||||
OY 241 TTTCACAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 300
    |||||
DB 263 TTTCACAGAAATGAAAAAGTTTGGGATACAGAGTTCTGCTCAGTTTATCGAGCTGAC 322
    |||||
OY 301 ACGAAATTTTGGCCCAATTTTATCGAGCATATGTTGGTTGGCAAAAGCCGCAATG 360
    |||||
DB 323 ACGAAATTTTGGCCCAATTTTATCGAGCATATGTTGGTTGGCAAAAGCCGCAATG 382
    |||||
OY 361 AGTATCCCAATATGATTAATTTATATGCTCCGCCGCTTGCATCAACGAAGATAC 420
    |||||
DB 383 AGTATCCCAATATGATTAATTTATATGCTCCGCCGCTTGCATCAACGAAGATAC 442
    |||||
OY 421 CAAAATGATATGTAATTTATGTAATATTCATGATGCGCAATTCGCGCAACT 480
    |||||
DB 443 CAAAATGATATGTAATTTATGTAATATTCATGATGCGCAATTCGCGCAACT 502
    |||||
OY 481 TCGCTGCTCCAAAATTCGATGAAGGAGCTCTCCTAAGCAAGCATCGTTCAAGTT 540
    |||||
DB 503 TCGCTGCTCCAAAATTCGATGAAGGAGCTCTCCTAAGCAAGCATCGTTCAAGTT 562
    |||||
OY 541 GGACAAGCTCTTGAACAT 559
    |||||
DB 563 GGACAAGCTCTTGAACAT 581
    |||||

RESULT 3
```


LOCUS	Bj115564	519 bp	mrna	linear	EST 23-Jan-2002
DEFINITION	Bj115564 unpublished oligo-capped cDNA library, C. elegans L1 stage				
ACCESSION	Caenorhabditis elegans cDNA clone yk1192h09 5', mRNA sequence.				
VERSION	Bj115564				
KEYWORDS	Bj115564.1 GI:18275658				
ORGANISM	EST.				
SOURCE	Caenorhabditis elegans				
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae				
AUTHORS	1 (bases 1 to 519)				
TITLE	Kohara, Y., Shih, I.-T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.				
JOURNAL	A complementary view of the C. elegans genome				
COMMENT	Unpublished				
FEATURES	Contact: Tadasu Shih-i				
SOURCE	Center for Genetic Resource Information				
	National Institute of Genetics				
	1111 Yata, Mishima, Shizuoka 411-8540, Japan				
	Tel: 81-559-81-6856				
	Fax: 81-559-81-6855				
	Email: tshini@genes.nig.ac.jp.				
	Location/Qualifiers				
	1. 519				
	/organism="Caenorhabditis elegans"				
	/mol_type="mRNA"				
	/strain="N2"				
	/db_xref="taxon:6239"				
	/clone="yk1192h09"				
	/sex="hermaphrodite"				
	/tissue_type="whole animal"				
	/dev_stage="L1"				
	/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"				
BASE COUNT	170 a	91 c	109 g	149 t	
ORIGIN					
Query Match	30.3%; Score 453; DB 12; Length 519;				
Best Local Similarity	99.8%; Pred. No. 1.9e-233;				
Matches	503; Conservative	0; Mismatches	1; Indels	0; Gaps	0;
QY	1	ATGTCGATTTCTGAAAAATTTGTGACAGCTAACAAAAATCGACAGAAAACTCGATTAAG	60		
DB	16	ATGTCGATTTCTGAAAAATTTGTGACAGCTAACAAAAATCGACAGAAAACTCGATTAAG	75		
QY	61	ACCTACTTGTGGGAATCCTATTATACATCAGTTCGAGAAAGAAAACCTCTTTCAATTC	120		
DB	76	ACCTACTTGTGGGAATCCTATTATACATCAGTTCGAGAAAGAAAACCTCTTTCAATTC	135		
QY	121	GTTGAGACATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTAAGAGAGGAGTT	180		
DB	136	GTTGAGACATTCATCGTAACCTTACAGTTAATTTTAAAGAAATCGTAAGAGAGGAGTT	195		
QY	181	ATCTTCGAACAGTGTCCATGATTTATGACAAAGAACTGCATTCAGATCAGATGG	240		
DB	196	ATCTTCGAACAGTGTCCATGATTTATGACAAAGAACTGCATTCAGATCAGATGG	255		
QY	241	TTTGCAGCAATTTGAAAAAGTTGGCGGATACAGAGTTCTGCTCAGATTATTCGAGCTGAC	300		
DB	256	TTTGCAGCAATTTGAAAAAGTTGGCGGATACAGAGTTCTGCTCAGATTATTCGAGCTGAC	315		
QY	301	ACGAAATTTTGGCCCAATATTTTATPNDGAGCATATGTTGGTTGGCAAGCCGCAATG	360		
DB	316	ACGAAATTTTGGCCCAATATTTTATPNDGAGCATATGTTGGTTGGCAAGCCGCAATG	375		
QY	361	AGTATCCCAATATGATGATAAATTTATATCTCCGCCGCTTGGCAATCAAGAGAAATAC	420		
DB	376	AGTATCCCAATATGATGATAAATTTATATCTCCGCCGCTTGGCAATCAAGAGAAATAC	435		
QY	421	CAAAATGATATGTPAATTTATGTAATTAATTCATGATGGCGAAATCGTCGGCCAAACT	480		
DB	436	CAAAATGATATGTPAATTTATGTAATTAATTCATGATGGCGAAATCGTCGGCCAAACT	495		

QY	481	TCGCTGTCCTCCAAAATTCGATGAA	504
DB	496	TCGCTGTCCTCCAAAATTCGATGAA	519
RESULT 4			
LOCUS	BJ124772	557 bp	mRNA linear
DEFINITION	BJ124772	unpublished oligo-capped cDNA library, C. elegans l1 stage	
ACCESSION	Bj124772	Caenorhabditis elegans CDNA clone YK1315b1 5', mRNA sequence.	
VERSION	Bj124772.2	GI:3124608	
KEYWORDS	EST.		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae; Rhabditidae; Pelodierinae; Caenorhabditis.		
AUTHORS	1 (bases 1 to 557) Kohara,Y., Shin-I,T., Thierry-Mieg,D., Thierry-Mieg,D., Suzuki,Y. and Sugano,S.		
TITLE	A complementary view of the C.elegans genome		
JOURNAL	Unpublished		
COMMENT	On Jan 23, 2002 this sequence version replaced gi:1828913. Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshini@genes.nig.ac.jp.		
FEATURES	location/Qualifiers		
source	1..557		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="YK1315b1"		
	/sex="hermaphrodite"		
	/tissue_type="whole animal"		
	/dev_stage="L1"		
	/clone_11b="unpublished oligo-capped cDNA library, C. elegans l1 stage"		
BASE COUNT	180 a 99 c 118 g 158 t 2 others		
ORIGIN			
Query Match	28.9%; Score 433; DB 12; Length 557;		
Best Local Similarity	100.0%; Pred. No. 1.3e-222;		
Matches	433; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	102	AAAAACTCTTCATTCACGTTGAAGCATTCATTCGTAACCTTACAGTTAATTTTAACGA	161
DB	125	AAAAACTCTTCATTCACGTTGAAGCATTCATTCGTAACCTTACAGTTAATTTTAACGA	184
QY	162	ATGCGTGAAGAGAGAGTATCTTCGGAACAGTGGTCATCATTAATGACAAAGACTGGCA	221
DB	185	ATGCGTGAAGAGAGAGTATCTTCGGAACAGTGGTCATCATTAATGACAAAGACTGGCA	244
QY	222	TTGCAATTCAGTCAGATGTTTGCACGAATTTGAAAAAGTTTGGGATACAGAGTTCTGGC	281
DB	245	TTGCAATTCAGTCAGATGTTTGCACGAATTTGAAAAAGTTTGGGATACAGAGTTCTGGC	304
QY	282	TCAGTTATTCGAGAGTGCACAGAAATTTTGGCTCAATATTTTATTCGAGCATATGTTGG	341
DB	305	TCAGTTATTCGAGAGTGCACAGAAATTTTGGCTCAATATTTTATTCGAGCATATGTTGG	364
QY	342	TTTGGAACAACCCCGCAATGAGTGAATCCCAATGAGTAAATTTGATATGCTCCGCGCT	401
DB	365	TTTGGAACAACCCCGCAATGAGTGAATCCCAATGAGTAAATTTGATATGCTCCGCGCT	424
QY	402	TGCAATCAACGAGAGATACCAAAATGATATGTTGTAATATGTAATAATTCATTTGATG	461
DB	425	TGCAATCAACGAGAGATACCAAAATGATATGTTGTAATATGTAATAATTCATTTGATG	484

QY 462 CGAATCGTCGCCAACAATTGCTGCTCCAAAATTCGATGAAGGAGGCTCTCTAG 521
|||||
DB 485 CGAATCGTCGCCAACAATTGCTGCTCCAAAATTCGATGAAGGAGGAGGCTCTCTAG 544
QY 522 CAAGCATGCTTTC 534
|||||
DB 545 CAAGCATGCTTTC 557

RESULT 5
BU143186/c 768 bp mRNA linear EST 23-JAN-2002
LOCUS BU143186 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1192h09 3', mRNA sequence.
ACCESSION BU143186
VERSION BU143186.1 GI:18303352
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Pelodertinae; Caenorhabditis.
AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.
FEATURES
source Location/Qualifiers
1..768
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1192h09"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C.
elegans L1 stage"
BASE COUNT 232 a 151 c 146 g 237 t 2 others
ORIGIN
Query Match 27.7%; Score 414; DB 12; Length 768;
Best Local Similarity 100.0%; Pred. No. 2,9e-212;
Matches 414; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1147 TTCAGAGACATTCAGATGGATGAATACTTGGAGAGAAATCTCGAAGACCTTACG 1206
|||||
DB 468 TTCAGAGACATTCAGATGGATGAATACTTGGAGAGAAATCTCGAAGACCTTACG 409
QY 1207 CTGACCTGTTCAAGCAATGCTCTCCAGAGAGATTAGACAATTTAAGTA 1260
|||||
DB 408 CTGACCTGTTCAAGCAATGCTCTCCAGAGAGATTAGACAATTTAAGTA 355

RESULT 6
CB404801/c 607 bp mRNA linear EST 15-MAY-2003
LOCUS CB404801 AD-wrmCDNA Caenorhabditis elegans cDNA, mRNA sequence.
DEFINITION OSTR028D3_1
ACCESSION CB404801
VERSION CB404801.1 GI:30746528
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Pelodertinae; Caenorhabditis.
AUTHORS Reboul, J., Vagillo, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong
, C.M., Li, S., Jacobot, L., Berlin, N., Janky, R., Moore, T., Hudson
, J.R., Hartley, J.L., Brasch, M.A., Vandenhaute, J., Boulton, S.,
Endress, G.A., Jenna, S., Chevet, E., Papanicolaou, V., Tollas, P.P.,
Placek, J., Snyder, M., Huang, R., Chance, M.R., Lee, H.,
Doucette-Stamm, L., Hill, D.E. and Vidal, M.
TITLE C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
JOURNAL Nat. Genet., (2003) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david.hill@dfci.harvard.edu or
marc.vidal@dfci.harvard.edu
FEATURES
source Location/Qualifiers
1..607
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="Hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmCDNA"
/note="The AD-wrmCDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pCR86"
BASE COUNT 175 a 131 c 113 g 188 t
ORIGIN
Query Match 25.7%; Score 385; DB 14; Length 607;
Best Local Similarity 100.0%; Pred. No. 1.4e-196;
Matches 385; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 996 CAAATTTTGCAGAACTGAAGATATCTTATTTGTGGAGATGATGTCAGATGCACATTGA 1055
|
|
|
Db 454 CAAATTTTGCAGAACTGAAGATATCTTATTTGTGGAGATGATGTCAGATGCACATTGA 395
|
|
|
QY 1056 AGCAGATTTTCCATTCATATCATATATACATTATTTGTTCCAGTTGTTATGCGGAAA 1115
|
|
|
Db 394 AGCAGATTTTCCATTCATATCATATATACATTATTTGTTCCAGTTGTTATGCGGAAA 335
|
|
|
QY 1116 GATATATTTGGAACCTGTTCCCGCCAGATGATGTCAGAAAGCAATTCAGATGGGATGAAT 1175
|
|
|
Db 334 GATATATTTGGAACCTGTTCCCGCCAGATGATGTCAGAAAGCAATTCAGATGGGATGAAT 275
|
|
|
QY 1176 CTTGGAGAAAGATTCGAGAAACCCCTACCGCTTGACTTTGTCAGACCAATGCTTCCCA 1235
|
|
|
Db 274 CTTGGAGAAAGATTCGAGAAACCCCTACCGCTTGACTTTGTCAGACCAATGCTTCCCA 215
|
|
|
QY 1236 AGAGAGATTAGACAAATTTAAGCTA 1260
|
|
|
Db 214 AGAGAGATTAGACAAATTTAAGCTA 190
|
|
|
RESULT 7
C42305 360 bp mRNA linear EST 18-OCT-1999
LOCUS C42305 Yui1 Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
DEFINITION Caenorhabditis elegans cDNA clone yk292d9 5', mRNA sequence.
ACCESSION C42305
KEYWORDS EST.
SOURCE EST.
ORGANISM Caenorhabditis elegans
          Caenorhabditis elegans
          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea
          1 (bases 1 to 360)
REFERENCE Kohara,Y., Motobashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
AUTHORS M., Miyata,A. and Nishigaki,A.
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yui1 Kohara
          Genome Biology Lab.
          National Institute of Genetics
          Yata 1111, Mishima, Shizuoka 411, Japan
          Tel: 81-559-81-6854
          Fax: 81-559-81-6855
          Email: ykohara@lab.nig.ac.jp.
          Location/Qualifiers
FEATURES
  source
    1..360
    /organism="Caenorhabditis elegans"
    /mol_type="mRNA"
    /strain="N2"
    /db_xref="taxon:6239"
    /clone="yk292d9"
    /sex="hermaphrodite"
    /dev_stage="embryo"
    /clone_lib="Yui1 Kohara unpublished cDNA:Strain N2
    hermaphrodite embryo"
BASE COUNT 119 a 62 c 78 g 101 t
ORIGIN
Query Match 24.0%; Score 360; DB 14; Length 360;
Best Local Similarity 100.0%; Pred. No. 4e-183;
Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db 121 TACAGTTAAATTTTACGAATGCGTGAAGAGGATTAATCTTCGAAACAGTGGCCATGA 180
|
|
|
QY 204 TTATGACAGAACTGGATTCGATTCAGTCAGATGTTGACGAATGAAGAAAGTTGG 263
|
|
|
Db 181 TTATGACAGAACTGGATTCGATTCAGTCAGATGTTGACGAATGAAGAAAGTTGG 240
|
|
|
QY 264 CGGATCAGAGTTCTGCTCAGTTTATCGAGCTGACAGAAATTTGGCTCAATATTTT 323
|
|
|
Db 241 CGGATCAGAGTTCTGCTCAGTTTATCGAGCTGACAGAAATTTGGCTCAATATTTT 300
|
|
|
QY 324 ATCGACGATATGTTGGTTGGCAAAAGCCGCAATGAGTATCCCAATATGATTAAT 383
|
|
|
Db 301 ATCGACGATATGTTGGTTGGCAAAAGCCGCAATGAGTATCCCAATATGATTAAT 360
|
|
|
RESULT 8
Bj129361 761 bp mRNA linear EST 23-JAN-2002
LOCUS Bj129361 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1029h03 3', mRNA sequence.
ACCESSION Bj129361
KEYWORDS EST.
SOURCE EST.
ORGANISM Caenorhabditis elegans
          Caenorhabditis elegans
          Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea
          1 (bases 1 to 761)
REFERENCE Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
AUTHORS and Sugano,S.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasi Shin-I
          Center For Genetic Resource Information
          National Institute of Genetics
          1111 Yata, Mishima, Shizuoka 411-8540, Japan
          Tel: 81-559-81-6856
          Fax: 81-559-81-6855
          Email: tshin@genes.nig.ac.jp.
          Location/Qualifiers
FEATURES
  source
    1..761
    /organism="Caenorhabditis elegans"
    /mol_type="mRNA"
    /strain="N2"
    /db_xref="taxon:6239"
    /clone="yk1029h03"
    /sex="hermaphrodite"
    /tissue_type="whole animal"
    /dev_stage="L1"
    /clone_lib="unpublished oligo-capped cDNA library, C.
    elegans L1 stage"
BASE COUNT 233 a 150 c 144 g 232 t 2 others
ORIGIN
Query Match 23.8%; Score 356; DB 12; Length 761;
Best Local Similarity 99.8%; Pred. No. 7.6e-181;
Matches 406; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      1094  TCCAGTTGGTTATGCGAAAGATATATTGGAACTGTTCGCCGAGATGAGTTCAAG 1153
          |||
          521  TCCAGTTGGTTATGCGAAAGATATATTGGAACTGTTCGCCGAGATGAGTTCAAG 462
          |||
QY      1154  GAACATTCACATGGGATGATCTTGGAGAAAGATTCGCAAAACCCCTACCGTTGACT 1213
          |||
          461  GAACATTCACATGGGATGATCTTGGAGAAAGATTCGCAAAACCCCTACCGTTGACT 402
          |||
QY      1214  TGTTCAGACCAATGCTTCCCAAGAGATTAGACAATTTAAGTA 1260
          |||
          401  TGTTCAGACCAATGCTTCCCAAGAGATTAGACAATTTAAGTA 355
          |||

RESULT 9
LOCUS   AU216908/c 677 bp mRNA linear EST 17-JUL-2001
DEFINITION AU216908 unpublished oligo-capped cDNA library, stage 1d
            Caenorhabditis elegans cDNA clone yk843e05 3', mRNA sequence.
ACCESSION AU216908
VERSION   AU216908.1 GI:14855065
KEYWORDS  EST.
SOURCE    Caenorhabditis elegans
ORGANISM  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea
REFERENCE 1 (bases 1 to 677)
AUTHORS   Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
            and Sugano,S.
            A complementary view of the C.elegans genome
TITLE      Unpublished
JOURNAL    Contact: Yuji Kohara
COMMENT    Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES   Location/Qualifiers
            source
            1..677
            /organism="Caenorhabditis elegans"
            /mol_type="mRNA"
            /strain="N2"
            /db_xref="taxon:6239"
            /clone="yk843e05"
            /sex="Hermaphrodite"
            /tissue_type="whole animal"
            /dev_stage="L4"
            /clone_lib="unpublished oligo-capped cDNA library, stage
            L4"

BASE COUNT 210 a 136 c 130 g 201 t

Query Match 22.3%; Score 334; DB 9; Length 677;
Best Local Similarity 100.0%; Pred. No. 6e-169;
Matches 334; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      927  TGACAAAGTTGAGCTCATGACCCCTTGCTCAGCAATTCATAACCTCCAGTGC 986
          |||
          677  TGACAAAGTTGAGCTCATGACCCCTTGCTCAGCAATTCATAACCTCCAGTGC 618
          |||
QY      987  TTGCATCTCAAAATTTTGCAAAAGTGAAGATATCTTTGTGGAAATGATGGTCCGA 1046
          |||
          617  TTGCATCTCAAAATTTTGCAAAAGTGAAGATATCTTTGTGGAAATGATGGTCCGA 558
          |||
QY      1047  TGCACTTGAAGACAGTTTCTTATTCATATCAATTAATCATTTATGTTCCAGTTGTTA 1106
          |||
          557  TGCACTTGAAGACAGTTTCTTATTCATATCAATTAATCATTTATGTTCCAGTTGTTA 498
          |||
QY      1107  TGGCGAAAGTATATTTGGAAGTGTTCGCCAGATGAGTTCAAGGACATTCAGATG 1166
          |||
          497  TGGCGAAAGTATATTTGGAAGTGTTCGCCAGATGAGTTCAAGGACATTCAGATG 438
          |||

```

```

QY      1167  GGATGAATTAATTTGGAGAAAGATCTGCAGAAACCTTACCGCTTGACTTGTTCAGGCCAT 1226
          |||
          437  GGATGAATTAATTTGGAGAAAGATCTGCAGAAACCTTACCGCTTGACTTGTTCAGGCCAT 378
          |||
QY      1227  GCCTTCCCAAGAGATTAGACAATTTAAGTA 1260
          |||
          377  GCCTTCCCAAGAGATTAGACAATTTAAGTA 344
          |||

```

```

RESULT 10
LOCUS   AV197399 380 bp mRNA linear EST 26-JUL-1999
DEFINITION AV197399 yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
            embryo Caenorhabditis elegans cDNA clone yk65h11 5', mRNA
            sequence.
ACCESSION AV197399
VERSION   AV197399.1 GI:5581170
KEYWORDS  EST.
SOURCE    Caenorhabditis elegans
ORGANISM  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditiida; Rhabditoidea
REFERENCE 1 (bases 1 to 380)
AUTHORS   Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,
            Nishigaki,A., Motobashi,T., Zeng,Q., Matanabe,H., Sugimoto,A., Sano
            M., Miyata,A., Mitani,Y., Iida,K., Uesugi,H., Sugiyama,Y. and
            Nomoto,H.
            Expressed genes in C.elegans
TITLE      Unpublished
JOURNAL    Contact: Yuji Kohara
COMMENT    Genome Biology Lab.
            National Institute of Genetics
            Yata 1111, Mishima, Shizuoka 411, Japan
            Tel: 81-559-81-6854
            Fax: 81-559-81-6855
            Email: ykohara@lab.nig.ac.jp.
FEATURES   Location/Qualifiers
            source
            1..380
            /organism="Caenorhabditis elegans"
            /mol_type="mRNA"
            /strain="N2"
            /db_xref="taxon:6239"
            /clone="yk65h11"
            /sex="hermaphrodite"
            /dev_stage="embryo"
            /clone_lib="yuji Kohara unpublished cDNA:Strain N2
            hermaphrodite embryo"

BASE COUNT 125 a 63 c 83 g 108 t 1 others
ORIGIN

```

```

Query Match 21.5%; Score 322; DB 9; Length 380;
Best Local Similarity 100.0%; Pred. No. 1.e-162;
Matches 322; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      71  GGAATCCTATTTATCATGATTCGAGAAAGAAACTTCTTTCATTCAGTTGAAGCAT 130
          |||
          59  GGAATCCTATTTATCATGATTCGAGAAAGAAACTTCTTTCATTCAGTTGAAGCAT 118
          |||
QY      131  TCAATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAAAGATTATCTTCGAA 190
          |||
          119  TCAATCGTAACCTTACAGTTAATTTTAAAGAAATCGTGAAGAAAGATTATCTTCGAA 178
          |||
QY      191  CAGTGGTCCATGATTTAGACAAGAACTGCGATTGATTCAGATCAAGTGGTTGACGAA 250
          |||
          179  CAGTGGTCCATGATTTAGACAAGAACTGCGATTGATTCAGATCAAGTGGTTGACGAA 238
          |||
QY      251  TTGAAGAAAGTTGGGATAGACAGTTCGTCAGTTATATGGAGCTGACAGAAATTTT 310
          |||
          239  TTGAAGAAAGTTGGGATAGACAGTTCGTCAGTTATATGGAGCTGACAGAAATTTT 298
          |||
QY      311  GGCCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAACGCCGCAATGAGATGCCA 370
          |||
          299  GGCCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAACGCCGCAATGAGATGCCA 358
          |||

```

OY 371 ATATGATAAATGTATATGC 392
 DB 359 ATATGATAAATGTATATGC 380

RESULT 11

LOCUS AU111588 340 bp mRNA linear EST 19-OCT-2000
 DEFINITION AU111588 unpublished oligo-capped cDNA library Caenorhabditis
 elegans cDNA clone yk732e5 5', mRNA sequence.

ACCESSION AU111588
 VERSION AU111588.1 GI:10925155
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 340)
 AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.

TITLE A complementary view of the C. elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Yuji Kohara
 Genome Biology Lab.
 National Institute of Genetics
 Yata 1111, Mishima, Shizuoka 411, Japan
 Tel: 81-559-81-6854
 Fax: 81-559-81-6855
 Email: ykohara@lab.nig.ac.jp.

FEATURES
 source
 1..340
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk732e5"
 /sex="Hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="varied"
 /clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 111 a 58 c 72 g 98 t 1 others
 ORIGIN

Query Match 20.6%; Score 308; DB 9; Length 340;
 Best Local Similarity 100.0%; Pred. No. 5,7e-155;
 Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATAGAC 62
 DB 33 GTCTGAATTTCTGAAATTTGTCAGAGCTAACAAAAAATCGACAGAAAACTCGATAGAC 92
 OY 63 CTACTGTGGGAATCTATTACATCAGTCGAGAAAGAAATCTTTCATTCAGT 122
 DB 93 CTACTGTGGGAATCTATTACATCAGTCGAGAAAGAAATCTTTCATTCAGT 152
 OY 123 TGAACATTCATCTTAACCTTACAGTTAAATTTAAACGAATCGTGAAGAGAGATTAT 182
 DB 123 TGAACATTCATCTTAACCTTACAGTTAAATTTAAACGAATCGTGAAGAGAGATTAT 212
 OY 183 CTTCGAAACAGTGTCCATGATTATGACAGAACTGCGATTCGATTCAAGTCAGATGTT 242
 DB 213 CTTCGAAACAGTGTCCATGATTATGACAGAACTGCGATTCGATTCAAGTCAGATGTT 272
 OY 243 TGCAGCAATGAAAGTTTGGCGATACAGAGTTGCTGCTTATTCGAGCTGACAC 302
 DB 273 TGCAGCAATGAAAGTTTGGCGATACAGAGTTGCTGCTTATTCGAGCTGACAC 332
 OY 303 GAAATTTT 310
 DB 333 GAAATTTT 340

RESULT 12
 LOCUS BJ153087/c 686 bp mRNA linear EST 24-JAN-2002
 DEFINITION BJ153087 unpublished oligo-capped cDNA library, C. elegans L1 stage
 Caenorhabditis elegans cDNA clone yk1315b11 3', mRNA sequence.

ACCESSION BJ153087
 VERSION BJ153087.1 GI:18321072
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
 ; Rhabditidae; Pelodierinae; Caenorhabditis.
 1 (bases 1 to 686)
 AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
 and Sugano,S.

TITLE A complementary view of the C. elegans genome
 JOURNAL Unpublished
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsuhiegenes.nig.ac.jp.

FEATURES
 source
 1..686
 Location/Qualifiers
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="yk1315b11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"

BASE COUNT 214 a 133 c 131 g 204 t 4 others
 ORIGIN

Query Match 20.4%; Score 306; DB 12; Length 686;
 Best Local Similarity 100.0%; Pred. No. 8.9e-154;
 Matches 306; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 955 TTGGCTCAGCAATTCATTAACCTCCAGCGCTGCTGATTCCTCAATTTTGGCAAACTGAA 1014
 DB 666 TTGGCTCAGCAATTCATTAACCTCCAGCGCTGCTGATTCCTCAATTTTGGCAAACTGAA 607
 OY 1015 GGATATCTTATTTGGGAATGAGTGTCCAGATGCACTTGAAGACAGTTTCTTATTCAT 1074
 DB 606 GGATATCTTATTTGGGAATGAGTGTCCAGATGCACTTGAAGACAGTTTCTTATTCAT 547
 OY 1075 ATCAATTAATACATTTATGTTCCAGTGTGTTGCGAAAGTATTAATTTGGACCTGTT 1134
 DB 546 ATCAATTAATACATTTATGTTCCAGTGTGTTGCGAAAGTATTAATTTGGACCTGTT 487
 OY 1135 CCGCCAGATGAGTTCAAGGAACATTCAGATGAGTGAATACTTGAGAAAGAAATCTGCA 1194
 DB 486 CCGCCAGATGAGTTCAAGGAACATTCAGATGAGTGAATACTTGAGAAAGAAATCTGCA 427
 OY 1195 GAAACCTTACCGCTTGACTTGTTCAGGCCAATGCTTCCCAAGAGAGATTAGACAAATTT 1254
 DB 426 GAAACCTTACCGCTTGACTTGTTCAGGCCAATGCTTCCCAAGAGAGATTAGACAAATTT 367
 OY 1255 AAGGTA 1260
 DB 366 AAGGTA 361

RESULT 13
 LOCUS AU116296/c 300 bp mRNA linear EST 19-OCT-2000
 DEFINITION AU116296 unpublished oligo-capped cDNA library Caenorhabditis
 elegans cDNA clone yk743a4 3', mRNA sequence.

ACCESSION	AU116296	GI:10929863	
VERSION	AU116296.1		
KEYWORDS	EST		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	1 (bases 1 to 300)		
	Kohara,Y., Shlin-1,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.		
	and Sugano,S.		
TITLE	A complementary view of the C. elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab. National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohara@lab.nig.ac.jp.		
FEATURES	location/Qualifiers		
source	1..300		
	/organism="Caenorhabditis elegans"		
	/mol_type="mRNA"		
	/strain="N2"		
	/db_xref="taxon:6239"		
	/clone="yk743a4"		
	/sex="Hermaphrodite"		
	/tissue.type="whole animal"		
	/dev_stage="varied"		
	/clone.lib="unpublished oligo-capped cDNA library"		
BASE COUNT	95 a 63 c 54 g 88 t		
ORIGIN			
Query Match	14.7%; Score 220; DB 9; Length 300;		
Best Local Similarity	100.0%; Pred. No. 2.5e-107;		
Matches	220; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
OY	1278 GGTAGAGCTACGCGCTTGAGAGCTGCTGCATGTGTGAAAATACAGTTATTGTCACAGCTAC	1337	
DB	300 GGTAGAGCTACGCGCTTGAGAGCTGCTGCATGTGTGAAAATACAGTTATTGTCACAGCTAC	241	
OY	1338 AGTGAATCATCTCATGAGAACATGATAAATGTCATATTGCGACGGCTGGATGAAGAATT	1397	
DB	240 AGTGAATCATCTCATGAGAACATGATAAATGTCATATTGCGACGGCTGGATGAAGAATT	181	
OY	1398 TGATGACTGTATGATGTGACATCCCATGATATTTCTACCGATAGAGTGTGAAGCGCA	1457	
DB	180 TGATGACTGTATGATGTGACATCCCATGATATTTCTACCGATAGAGTGTGAAGCGCA	121	
OY	1458 CAGTTATGTTACACACTCCGAAAAGATACACATATGA	1497	
DB	120 CAGTTATGTTACACACTCCGAAAAGATACACATATGA	81	
RESULT 14			
AU115428/c	300 bp	mrna	linear
LOCUS	AU115428	unpublished oligo-capped cDNA library	EST 19-OCT-2000
DEFINITION	AU115428	elegans cDNA clone yk7325 3', mRNA sequence.	
ACCESSION	AU115428		
VERSION	AU115428.1	GI:10928995	
KEYWORDS	EST		
SOURCE	Caenorhabditis elegans		
ORGANISM	Caenorhabditis elegans		
REFERENCE	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea		
AUTHORS	1 (bases 1 to 300)		
	Kohara,Y., Shlin-1,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.		
	and Sugano,S.		
TITLE	A complementary view of the C. elegans genome		
JOURNAL	Unpublished		
COMMENT	Contact: Yuji Kohara Genome Biology Lab.		

FEATURES		source		National Institute of Genetics Yata 1111, Mishima, Shizuoka 411, Japan Tel: 81-559-81-6854 Fax: 81-559-81-6855 Email: ykohnara@lab.nig.ac.jp. Location/Qualifiers	
FEATURES		source		1. .300 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="yk732e5" /sex="Hermaphrodite" /tissue_type="whole animal" /dev_stage="varied" /clone.lib="unpublished oligo-capped cDNA library"	
BASE COUNT		94 a 60 c 54 g 89 t 3 others			
ORIGIN					
Query Match		14.4%; Score 215; DB 9; Length 300;			
Best Local Similarity		100.0%; Pred. No. 1.3e-104;			
Matches 215; Conservative		0; Mismatches 0; Indels 0; Gaps 0;			
QY	1283	GACTACGCTTGAAAGCTGCTCACAATGTGTGAAAAATCACTTATTATTTGTCACACTCACTGA	1342		
Db	300	GACACAGCGCTTGAAGCTGCTCACAATGTGTGAAAAATCACTTATTATTTGTCACACTCACTGA	241		
QY	1343	AATCAGTTCATGAGAAGACTGATTAATATGCAATTTTCAGACGCGTGGGATGAAGATTTGATG	1402		
Db	240	AATCAGTTCATGAGAAGACTGATTAATATGCAATTTTCAGACGCGTGGGATGAAGATTTGATG	181		
QY	1403	AACTGTATGATGTGCACTCCCATGATATTTCACGATAGATGTGTGAACGCACAGTT	1462		
Db	180	AACTGTATGATGTGCACTCCCATGATATTTCACGATAGATGTGTGAACGCACAGTT	121		
QY	1463	ATGTTCTACACCTCCGAAAAAGTACACACTATTGA	1497		
Db	120	ATGTTCTACACCTCCGAAAAAGTACACACTATTGA	86		
RESULT 15					
LOCUS		A0112450		337 bp mRNA linear EST 30-MAY-2003	
DEFINITION		A0112450 unpublished oligo-capped cDNA library Caenorhabditis			
ACCESSION		A0112450			
VERSION		A0112450.2		GI:31237498	
KEYWORDS		EST.			
SOURCE		Caenorhabditis elegans			
ORGANISM		Caenorhabditis elegans			
REFERENCE		Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea			
AUTHORS		; Rhabditidae; Pelodermidae; Caenorhabditis.			
TITLE		1 (bases 1 to 337)			
JOURNAL		Kohnara,Y., Shih,I.T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.			
COMMENT		A complementary view of the C. elegans genome			
		Unpublished			
		On Oct 19, 2000 this sequence version replaced gi:10926017.			
		Contact: Yuji Kohnara			
		Genome Biology Lab.			
		National Institute of Genetics			
		Yata 1111, Mishima, Shizuoka 411, Japan			
		Tel: 81-559-81-6854			
		Fax: 81-559-81-6855			
		Email: ykohnara@lab.nig.ac.jp.			
FEATURES		Location/Qualifiers			
source		1. .337 /organism="Caenorhabditis elegans" /mol_type="mRNA" /strain="N2" /db_xref="taxon:6239" /clone="yk743a4" /sex="Hermaphrodite"			

/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"
BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN

Query Match 12.4%; Score 186; DB 9; Length 337;
Best Local Similarity 100.0%; Pred. No. 6.9e-89;
Matches 186; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGTCGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAACTCGATAG 60
|||
Db 23 ATGTCGAATTTCTGAAAAATGTCAGAGCTAACAAAAATCGACAGAAACTCGATAG 82
|||
QY 61 ACCTACTGTGGGAATCCATTACATCACTTCGAGAAAGGAAAACTTCTTCATTCCA 120
|||
Db 83 ACCTACTGTGGGAATCCATTACATCACTTCGAGAAAGGAAAACTTCTTCATTCCA 142
|||
QY 121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTACGAAATCGTGAAGAGAGTT 180
|||
Db 143 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTACGAAATCGTGAAGAGAGTT 202
|||
QY 181 ATCTTC 186
|||
Db 203 ATCTTC 208

Search completed: September 2, 2003, 20:31:03
Job time : 3446 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 16:44:20 ; Search time 5682 Seconds

(without alignments)
10778.196 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497

Sequence: 1 atgctgaattctgaaat.....cgaaaagtcacactatgta 1497

Scoring table: OLIGO_NUC

Searched: 2888711 seqs, 2045481386 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: listing first 45 summaries

Database : GenBankl:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fan:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	833	55.6	30911	3	CER06C7	271266 Caenorhabdi
C 2	25	1.7	218764	2	AC103229	AC103229 Rattus no
C 3	23	1.5	133343	2	AP005398	AP005398 Oryza sat
C 4	22	1.5	157654	2	AC091791	AC091791 Sus scrofi
C 5	22	1.5	183239	2	AC128697	AC128697 Mus muscu
C 6	22	1.5	193311	2	AC118694	AC118694 Mus muscu
C 7	22	1.5	195037	2	AC091790	AC091790 Sus scrofi
C 8	22	1.5	205682	2	AC140183	AC140183 Mus muscu
C 9	22	1.5	229331	2	AC118916	AC118916 Rattus no
C 10	22	1.5	231442	2	AC094981	AC094981 Rattus no
C 11	22	1.5	253961	2	AC128235	AC128235 Rattus no
C 12	21	1.4	872	6	BD146582	BD146582 Primer fo
C 13	21	1.4	2202	9	BD159393	BD159393 Primer fo
C 14	21	1.4	2202	9	AK022163	AK022163 Homo sapi
C 15	21	1.4	11922	3	DMC137E7	DMC137E7 Homo sapi
C 16	21	1.4	42052	3	U80029	U80029 Caenorhabdi
C 17	21	1.4	59824	2	AC109340	AC109340 Homo sapi
C 18	21	1.4	110000	2	AC105643	AC105643 Continuation (5 of
C 19	21	1.4	135580	2	AC017687	AC017687 Drosophill
C 20	21	1.4	140658	5	AF112374	AF112374 Dantio rer
C 21	21	1.4	146074	2	AC141806	AC141806 Apis mell
C 22	21	1.4	152575	2	AC141690	AC141690 Apis mell
C 23	21	1.4	166231	2	AC120880	AC120880 Mus muscu
C 24	21	1.4	170490	3	AC104147	AC104147 Drosophill
C 25	21	1.4	173525	2	AC067893	AC067893 Homo sapi
C 26	21	1.4	175262	3	AC098576	AC098576 Drosophill
C 27	21	1.4	176709	9	AC011737	AC011737 Homo sapi
C 28	21	1.4	184585	9	AC079465	AC079465 Homo sapi
C 29	21	1.4	184649	2	AC025185	AC025185 Homo sapi
C 30	21	1.4	187003	2	AC023353	AC023353 Homo sapi
C 31	21	1.4	196745	9	AC087431	AC087431 Homo sapi
C 32	21	1.4	204478	2	AC130948	AC130948 Rattus no
C 33	21	1.4	205066	2	AC120803	AC120803 Rattus no
C 34	21	1.4	220622	2	AC118570	AC118570 Lemur cat
C 35	21	1.4	240918	2	AC137254	AC137254 Rattus no
C 36	21	1.4	243833	2	AC097386	AC097386 Rattus no
C 37	21	1.4	251237	2	AC105634	AC105634 Rattus no
C 38	21	1.4	261251	2	AC096413	AC096413 Rattus no
C 39	21	1.4	261603	3	AC098318	AC098318 Rattus no
C 40	21	1.4	300933	3	AE003422	AE003422 Drosophill
C 41	21	1.4	323991	2	AC098512	AC098512 Rattus no
C 42	20	1.3	2044	6	AR302542	AR302542 Sequence
C 43	20	1.3	2863	6	AR302510	AR302510 Sequence
C 44	20	1.3	2863	6	AX058963	AX058963 Sequence
C 45	20	1.3	3646	6	AR302540	AR302540 Sequence

ALIGNMENTS

RESULT 1

CER06C7 30911 bp DNA linear INV 21-MAY-2003

LOCUS CER06C7/c

DEFINITION Caenorhabditis elegans cosmid R06C7, complete sequence.

ACCESSION Z71266

VERSION Z71266.1 GI:1279324

KEYWORDS HTG; Adenylosuccinate lyase; Myosin heavy chain; Rat TOAD-64

SOURCE Protein like; Serine/threonine-protein kinase; Zinc finger protein.

ORGANISM Caenorhabditis elegans

REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;

AUTHORS Rhabditoidea; Rhabditidae; Peloderiinae; Caenorhabditis.

TITLE none.

Pred. No. is the number of results predicted by chance to have a

JOURNAL Investigating biology. The C. elegans Sequencing Consortium
MEDLINE Science 282 (5396), 2012-2018 (1998)
99069613
REMARK The C.elegans Sequencing Consortium.
AUTHORS 2 (bases 1 to 30911)
TITLE Gardner,A.E.
JOURNAL Direct Submission

COMMENT

Submitted (19-APR-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
jes@sanger.ac.uk or rwhematode.wustl.edu
Coding sequences below are predicted from computer analysis, using
predictions from GeneFinder (P. Green, U. Washington), and other
available information.
Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is not the entire insert of clone R06C7.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone R06C7 is at 1 in this sequence. The true
right end of clone R06C7 is at 8718 in
sequence Z71261.
The true left end of clone F21C3 is at 30808 in this sequence. The
start of this sequence (1..106) overlaps with the end of sequence
Z98261.
The end of this sequence (30808..30911) overlaps with the start of
sequence Z71261.
For a graphical representation of this sequence and its analysis
see: - [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=R06C7)
name=R06C7
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
Location/Qualifiers
1..30911

FEATURES

source

/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="1"
/clone="R06C7"
join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))
/gene="C01H6.9"

gene

CDS

join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))
/gene="C01H6.9"
/standard_name="C01H6.9"
/note="contains similarity to Pfam domain: PF00069

(Eukaryotic protein kinase domain), Score=37.5,
E-value=8.2e-10, N=2
cDNA EST yk42h5.3 comes from this gene
cDNA EST yk42h5.5 comes from this gene"

/codon_start=1
/product="Hypothetical protein C01H6.9"

/protein_id="CAA85847.1"
/db_xref="GI:3878842"

/db_xref="GOA:O17985"
/db_xref="SPTRMBL:017985"

/translation="MPKPRIKVPISGKKARNFARNRNLSSVITTECPIDPEPE
FSKPALEKSPNDPEPKKKPKKPKPPPPSPDVSISDEENAGOKPDADA
DEEKIKIKIEADPSLDVSLSPYNDGIRKIKENGHIITHDKOPREAMPREPKITD
FAKMEOROKKPVKQVEPMNSKANQVCPNVAFCGSIADHRKKKSLAPSLF
RVKGTPORSKMIYPAKHVQCODERLAIYSTPIRVNNTAOTDITISIVSPKTONA
STPIEKOTAGRPKSPFPVPLNSDCENTKEEQJAOKSANSQDKETITNNPIE
EYSGRVESLSLAGSIQOPODLEETAISENKLSITKTPIESRNSHONKIMTOO
VSDSEVEDYETFNPKRFTSVTELLHEDEHEDSDNLFSKSKKROKRPSPKMO
GMSVWIMQDDDDLEALFEIEKNEIRKRTTLOPOOROPSSRSDSINSAMEKSLQO
FLEDTWEEFGKQYSESRAESRNIPTGMTIHNDPSILPFLYEDLVSIPSPMOL
LHVGOEKSTWDSLPKSAIDGRVKKLGGAAGEVSTIMDGPAVKITVPEKDCG
NROFGYEHSEEMQTSVDVLEPVYVMKELALRDEDMNSTPNFIEIMISAWEVKP
KGLSAMDYDKLKESENTPRDVYSIDONFLEVSANGIALDEDFYLSSENELFTII
HOLVLSMNAEALFEFHRDLHGIVNLIDNGYELVYTHGOKVPLSTGIKVNIID
FTLSRISKATVYWDLENDPAIFEGDDPOFVYRMRNCKSNMKKRSRRITLMTI
VYIANRLIDTRICPKGLITERKRMELVLDKRAEFSCGCSL7NEFPDFYEGPIG
MSTYRQ"

gene

CDS

complement(join(1315..1480,1528..1715,1747..1783,
1834..2156))
/gene="R06C7.2"
complement(join(1315..1480,1528..1715,1747..1783,
1834..2156))
/gene="R06C7.2"
/standard_name="R06C7.2"
/codon_start=1
/product="Hypothetical protein R06C7.2"
/protein_id="CAA95841.1"
/db_xref="GI:3878836"
/db_xref="SPTRMBL:Q21772"
/translation="MFSAVHPDSRRDCKKRYQSOTSHONYWFLKYSENNLGRKIKG
NFTLIFHAERKVLQSVYMLKQDKTKTRQETQCLANFENNRRKLQDTERN
KKLRNLINSELEKLVSSPENLPLOILFELISENDELERHFHCATFTLSDFOLF
FPPMPVYANYSKSGRLMWNPAITADIPVPSGEARMAQIOTDSIVKXSGFLIDBR
VMSALITQNSCSFENSNP"
join(3063..3227,3274..3854,3899..4631,4680..4855,
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
/gene="R06C7.1"
join(3063..3227,3274..3854,3899..4631,4680..4855,
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
/gene="R06C7.1"
/standard_name="R06C7.1"
/note="contains similarity to Pfam domain: PF02170 (2AP
domain), Score=151.7, E-value=4.2e-42, N=1; PF02171 (Plw1
domain), Score=456.7, E-value=6.3e-134, N=1
cDNA EST yk31a12.5 comes from this gene
cDNA EST yk21g1.3 comes from this gene
cDNA EST yk31a12.3 comes from this gene
cDNA EST yk31a12.3 comes from this gene
cDNA EST yk36g4.3 comes from this gene
cDNA EST yk31a4.3 comes from this gene
cDNA EST yk21g1.5 comes from this gene
cDNA EST yk21g1.5 comes from this gene
cDNA EST yk25b2.5 comes from this gene
cDNA EST yk34a6.5 comes from this gene
cDNA EST yk36g4.5 comes from this gene
cDNA EST yk31a4.5 comes from this gene
cDNA EST yk93h11.5 comes from this gene
cDNA EST yk125b7.5 comes from this gene
cDNA EST yk125b7.3 comes from this gene
cDNA EST yk236e4.3 comes from this gene
cDNA EST yk295d11.3 comes from this gene
cDNA EST yk348b1.3 comes from this gene
cDNA EST yk367e6.3 comes from this gene

```

CDNA EST yk236e4.5 comes from this gene
CDNA EST yk295d11.5 comes from this gene
CDNA EST yk348b1.5 comes from this gene
CDNA EST yk367e6.5 comes from this gene
CDNA EST yk313d8.3 comes from this gene
CDNA EST yk669d4.3 comes from this gene
CDNA EST yk671a8.3 comes from this gene
CDNA EST yk250f12.5 comes from this gene
CDNA EST yk513b8.5 comes from this gene
CDNA EST yk669p4.5 comes from this gene
CDNA EST yk671a8.5 comes from this gene
CDNA EST yk775g12.3 comes from this gene
CDNA EST yk775g12.5 comes from this gene
/product="Hypothetical protein R06c7.1"
/protein_id="CAA95839.1"
/db_xref="GI:3878834"
/db_xref="SPTREMBL:Q21770"
/translation="MSPHPPQHPMPMPVPTAPPGANTPMPVPADAKLHQTGN
DACLRLQDLNVEDGAKMYKPTPEKMKRPVDIOTNVGIEVTEKTYVHREMVAKA
DLSTREKVFTEKKEDEYVODRDRDCCIFLAEKNEPEFKMDGNITVDSGSTL
YTTVNLFSFLDANGKSKYEQINGADGDDDKLPCISLEIYARNDSNTISSSENGS
KRTADQNTIEVNNREYTOFLELALNQHCAVEINFGCFEFGKTYFLNATBEGDQRCV
DVGDGKOLYPGKTKTIOFTEGPGRQNNPSLVLDGKMAFKEDQIVLOKLDITGOD
PENGNNMTREKAAVYIKGLDYSTYNNKRRLRIEGLFHEBASTIRBELPDGKCSI
AEVYADKYKISLOYPNANLVCCDRGNMYPAEIMTVSRNORVITPOOTGOSOKT
KECAVLPDVRHMITGKNAVITLENELVALGIKVSSEPLMVAORELDGKELYOR
SYMDSMGKWRAPGPFVPRATPDYDLMAAVANGPCRSIGVNOVLGAFIDSCCKKG
MYIKRPPCTGLYSTERIKITOLEKVAASCKYVLMTTDDIVLHKYKALEORTMAYI
ODMKISKANAIVKDKRLTLENIINKTNVKGGLNTVSDARKSMTEQLIIGVGVSA
PRACTRYKMDNKNLPQIIGFASNAVAVHFEVGDVFLAPSGDPTMASIEDVLONSID
LEKRNKALPKRIIIRYSGASESHASIIAYEIPLARAIHGYSEIKLFIYVTRKH
SYRFRPDOLRSRGKATENMIPGIVDNTATPACQOPFNHTTLOGTAKTPIYVL
ADDCAKPMRELETFCTCHHOIYVSIPTPLLYVANEVAKRGDMLGELTTGPI
EAKESQGERLKELEIKGKQDMLNKRVA"
gene
complement(6739..11408)
/gene="dhp-1"
CDS
complement(Join(6739..6889,7899..7963,8094..8340,
8392..8457,8905..9182,9550..9761,9840..10017,10603..10783,
11317..11408))
/gene="dhp-1"

Query Match      55.6%; Score 833; DB 3; Length 30911;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      868 ACATTGATCAATTAGCAAAAGATCCATTGATCCCATGATTTGGAGAAAGTTAAGCTT 927
      17380 ACATTGATCAATTAGCAAAAGATCCATTGATCCCATGATTTGGAGAAAGTTAAGCTT 17321
Db
OY      928 GGACAAAAGTTGAGCTATGACACCCCTGGCTCAGCAATTCATTAACCTCGACGTGCT 987
      17320 GGACAAAAGTTGAGCTATGACACCCCTGGCTCAGCAATTCATTAACCTCGACGTGCT 17261
Db
OY      988 TCGATTCTCAAAATTTTGGCAAAACCTGAAGATATCTTATGTGGGATGGATGGTCCAGAT 1047
      17260 TCGATTCTCAAAATTTTGGCAAAACCTGAAGATATCTTATGTGGGATGGATGGTCCAGAT 17201
Db
OY      1048 GCACCTGAAGACAGTTTCTTATTCATTAATACATTAATATGTTTCCAGTGGTAT 1107
      17200 GCACCTGAAGACAGTTTCTTATTCATTAATACATTAATATGTTTCCAGTGGTAT 17141
Db
OY      1108 GCGGAAAAGTATTAATTTGGAACCTGTCCGCCAGATGAGTCAAGAAACATTCAGATGG 1167
      17140 GCGGAAAAGTATTAATTTGGAACCTGTCCGCCAGATGAGTCAAGAAACATTCAGATGG 17081
Db
OY      1168 GATGATCTTGGAGAAACATCTGAGAAACCTACCGCTGACTGTTCAGAACCAATG 1227
      17080 GATGATCTTGGAGAAACATCTGAGAAACCTACCGCTGACTGTTCAGAACCAATG 17021
Db
OY      1228 CCTCCCAAGAGATTTAGCAAAATTTAAGTAAATCTGATTTCCAAACGGGT 1280
      17020 CCTCCCAAGAGATTTAGCAAAATTTAAGTAAATCTGATTTCCAAACGGGT 16968
Db

RESULT 2
AC103229/C
LOCUS      Rattus norvegicus clone CH230-217018, WORKING DRAFT SEQUENCE.
DEFINITION AC103229
ACCESSION AC103229
VERSION    GI:30578703
KEYWORDS   HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
            Rattus.
            1 (bases 1 to 218764)
REFERENCE 1
AUTHORS   Muzny,D.,Marie,,Metzker,M.,Lee,,Abramzon,S.,Adams,C.,Alder,J.,
            Allen,C.,Allen,H.,Alsbrooks,S.,Amini,A.,Anguiano,D.,
            Anyalebechi,V.,Aoyagi,A.,Ayodeji,M.,Baca,E.,Baden,H.,
            Baldwin,D.,Bandaranaike,D.,Barber,M.,Barnstead,M.,Benahmed,F.,
            Biswalo,K.,Blair,J.,Blankenburg,K.,Blyth,P.,Brown,M.,
            Bryant,N.,Buhay,C.,Burch,P.,Burrell,K.,Calderton,E.,
            Cardenas,V.,Carter,K.,Cavazos,I.,Ceasar,H.,Center,A.,
            Chacko,J.,Chavez,D.,Chen,G.,Chen,R.,Chen,Y.,Chen,Z.,Chu,J.,
            Cleveland,C.,Cockrell,R.,Cox,C.,Coyle,M.,Cree,A.,D'Souza,L.,
            Davilla,M.L.,Davis,C.,Davy-Carroll,L.,De Anda,C.,Dexterich,D.,
            Delgado,O.,Denson,S.,Deramo,C.,Ding,Y.,Dinh,H.,Diyva,K.,
            Draper,H.,Dugan-Rocha,S.,Dunn,A.,Durbin,K.,Duvall,B.,Eaves,K.,
            Egan,A.,Escotto,M.,Eugene,C.,Evans,C.A.,Falls,T.,Fan,G.,
            Fernandez,S.,Finley,M.,Flagg,N.,Forbes,L.,Foster,M.,Foster,P.,
            Fraser,C.M.,Gabisel,A.,Ganta,R.,Garcia,A.,Garner,T.,Garza,M.,
            Gebregeorgis,E.,Geer,K.,Gilli,R.,Grady,M.,Guerra,M.,Guevara,W.,
            Gunaratne,P.,Haaland,W.,Hamill,C.,Hamilton,C.,Hamilton,J.,
            Harvey,Y.,Havlak,P.,Hawes,A.,Henderson,N.,Hernandez,J.,
            Hernandez,R.,Hines,S.,Hladun,S.L.,Hodgson,A.,Hogues,M.,
            Hollins,B.,Howells,S.,Hulk,S.,Hume,J.,Idlebird,D.,Jackson,A.,
            Jackson,L.,Jacob,L.,Jiang,H.,Johnson,B.,Johnson,R.,Jolyet,A.,
            Karpis,C.,Kelly,S.,Kelly,S.,Khan,Z.,King,L.,Kover,C.,
            Kowals,C.,Kraft,C.L.,Ledow,H.,Levan,J.,Lewis,L.,Li,Z.,Liu,J.,
            Liu,J.,Liu,W.,Liu,Y.,London,P.,Longacre,S.,Lopez,J.,
            Lorensuhewa,L.,Loulesed,H.,Lozado,R.J.,Lu,X.,Ma,J.,
            Maheshwari,M.,Mahindaratne,M.,Mahmoud,M.,Mailloy,K.,Mangum,A.,
            Mangum,B.,Mapua,P.,Martin,K.,McNeill,T.Z.,Meenen,E.,
            Mawhney,S.,McLeod,M.P.,McNeill,T.Z.,Meenen,E.,
            Milosavljevic,A.,Miner,G.,Ming,A.,Montemayor,J.,Moore,S.,
            Morgan,M.,Morris,K.,Morris,S.,Munidasa,M.,Murphy,M.,Nair,L.,

```

* NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

```

FEATURES
  source
    1. 218764
      Location/Qualifiers
        1. 218764
          /organism="Rattus norvegicus"
          /mol_type="genomic DNA"
          /db_xref="taxon:10116"
          /clone="CH230-217018"
          1. 1339
            /note="wgs_contig"

BASE COUNT  63585 a 42997 c 43704 g 64011 t 4467 others
ORIGIN
  misc_feature

Query Match
  Best Local Similarity 100.0%; Score 25; DB 2; Length 218764;
  Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY  5 CTGAATTTCTGAATAATGTCAGACG 29
    |||||||||||||||||||
Db  74190 CTGAATTTCTGAATAATGTCAGACG 74166

RESULT 3
AP005398/c 133343 bp DNA linear HTG 07-JUN-2002
LOCUS
DEFINITION
  Oryza sativa (japonica cultivar-group) chromosome 2 clone P0669G10,
  ** SEQUENCING IN PROGRESS **
ACCESSION
  AP005398
  AP005398.1 GI:21328228
KEYWORDS
  HTG; HTGS_PHASE2.
SOURCE
  Oryza sativa (japonica cultivar-group)
  Oryza sativa (japonica cultivar-group)
  Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
  1
  Sasaki,T., Matsumoto,T. and Katayose,Y.
  Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, PAC
  clone:P0669G10
  Published only in Database (2002)
  2 (bases 1 to 133343)
  Sasaki,T., Matsumoto,T. and Katayose,Y.
  Direct Submission
  Submitted (06-JUN-2002) Takuji Sasaki, National Institute of
  Agricultural Sciences, Rice Genome Research Program; Kannondai
  2-1-2, Tsukuba, Ibaraki 305-8602, Japan
  (E-mail:tsasaki@nris.affrc.go.jp; URL:http://rgrp.dna.affrc.go.jp/,
  Tel:81-298-38-7441, Fax:81-298-38-7468)
  NOTE: It currently consists of 1 contigs. Gaps between the contigs
  are represented as runs of N. The order of the pieces is believed
  to be correct as given, however the sizes of the gaps between them
  are based on estimates that have provided by the submitter. This
  sequence will be replaced by the finished sequence as soon as it is
  available and the accession number will be preserved.
  * NOTE: This is a 'working draft' sequence.
  * This sequence will be replaced
  * by the finished sequence as soon as it is available and
  * the accession number will be preserved.
  Location/Qualifiers
    1. 133343
      /organism="Oryza sativa (japonica cultivar-group)"
      /mol_type="genomic DNA"
      /cultivar="Nipponbare"
      /db_xref="taxon:39947"
      /chromosome="2"
      /clone="P0669G10"

BASE COUNT  38280 a 26701 c 27851 g 40458 t 53 others
ORIGIN
  Query Match
    1.5%; Score 23; DB 2; Length 133343;
  Best Local Similarity 100.0%; Pred. No. 0.92;
  Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

824 CAAATAAAATGAGAAATATCA 846

```


JOURNAL Submitted (25-AUG-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Aug 25, 2002 this sequence version replaced g1:21914538.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Project name: M.B80388P09
Center project name: M.B80388P09

----- Summary Statistics -----
Sequencing vector: M13, 08
Sequencing vector: plasmid: 1008
Chemistry: Dye-terminator ET: 0% of reads
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990319
Consensus quality: 177227 bases at least Q40
Consensus quality: 176348 bases at least Q30
Consensus quality: 176997 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 9 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

* 1 2586: contig of 2586 bp in length
* 2687 2686: gap of unknown length
* 2687 5769: contig of 3083 bp in length
* 5770 5869: gap of unknown length
* 5870 12437: contig of 6568 bp in length
* 12438 12537: gap of unknown length
* 12538 34518: contig of 21981 bp in length
* 34519 34618: gap of unknown length
* 34619 53625: contig of 19008 bp in length
* 53627 53726: gap of unknown length
* 53727 75556: contig of 21830 bp in length
* 75557 100967: gap of unknown length
* 100968 101067: gap of unknown length
* 101068 131088: contig of 30021 bp in length
* 131089 131188: gap of unknown length
* 131189 183239: contig of 52051 bp in length.

FEATURES
source
1. 183239
Location/Qualifiers

/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="X"
/clone="RP24-388P9"

misc_feature
/note="assembly_name:Contig9"
2687..5769
misc_feature
/note="assembly_name:Contig10"
5870..12437
misc_feature
/note="assembly_name:Contig11"
12538..34518
misc_feature
/note="assembly_name:Contig12"
34619..53626
misc_feature
/note="assembly_name:Contig13"
53727..75556
misc_feature
/note="assembly_name:Contig14"
75557..100967
misc_feature
/note="assembly_name:Contig15"
101068..131088
misc_feature
/note="assembly_name:Contig16"
131189..183239
misc_feature
/note="assembly_name:Contig17"
183239..52051
BASE COUNT 51407 a 40699 c 39395 g 50657 t 881 others

ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 183239;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 214 AACTGCATTCGATTCAGTCA 235
Db 153697 AACTGCATTCGATTCAGTCA 153676

RESULT 6

LOCUS AC118694 193311 bp DNA linear HTG 23-APR-2003
DEFINITION Mus musculus clone RP24-406H23, WORKING DRAFT SEQUENCE, 6 unordered
pieces.

ACCESSION AC118694
VERSION AC118694.6 GI:30018084
KEYWORDS HTG; HTGS; PHASE1; HTGS; DRAFT.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus

REFERENCE

AUTHORS

REFERENCE
TITLE
JOURNAL
REFERENCE
AUTHORS
Unpublished
2 (bases 1 to 193311)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N.,
Anderson, S., Brown, A., Cantarel, J., Chang, J.,
Boukhalter, B., Brown, A., Camarata, J., Campolongo, M., Collins, S., Collymore, A.,
Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Lander, T., Lehotzky, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., MacDonald, P., Major, J., Marquis, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L.,
McCarthy, T., Menega, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupack, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE

JOURNAL
REFERENCE
AUTHORS
Direct Submission
Submitted (20-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 193311)

Birren, B., Nusbaum, C., Lander, E., Abouelell, A., Allen, N.,
Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T.,
Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Choquet, Y.,
Collymore, A., Cook, A., Cooke, P., Corum, B., DeArrelano, K.,
Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S.,
Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S.,
Graham, L., Grand-pierre, N., Hafez, N., Hagopian, D., Hagos, B.,
Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R.,
Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C.,
MacDonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M.,
Meldrum, J., Menus, L., Mihova, T., Menega, V., Murphy, T., Naylor, J.,
Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P.,
O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N.,
Rachupka, A., Ramasamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C.,
Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M.,
Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M.,
Vassiliev, H., Venkataraman, V.S., Viel, R., Vo, A., Wilson, B., Wu, X.,

TITLE
JOURNAL
COMMENT

Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-APR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 17, 2003 this sequence version replaced gi:28467223.
All repeats were identified using RepeatMasker:
Smit,A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
----- Project Information
Center project name: L21128
Center clone name: 406_H23

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4849: contig of 4849 bp in length
* 4850 4949: gap of 100 bp
* 4950 99766: contig of 94817 bp in length
* 99867 99866: gap of 100 bp
* 99867 135133: contig of 35267 bp in length
* 135134 135233: gap of 100 bp
* 135234 156045: contig of 20812 bp in length
* 156046 156145: gap of 100 bp
* 156146 184734: contig of 28589 bp in length
* 184735 184834: gap of 100 bp
* 184835 193311: contig of 8477 bp in length.

FEATURES
source
1. 193311
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/clone="RP24-406H23"
/clone_1lb="RPC1-24 Male Mouse BAC"
1. 4849
/note="assembly-fragment
clone_end:SP6
vector_side:left"
4950..99766
/note="assembly-fragment"
99867..135133
/note="assembly-fragment"
135234..156045
/note="assembly-fragment"
156146..184734
/note="assembly-fragment"
184835..193311
/note="assembly-fragment
clone_end:T7
vector_side:right"

BASE COUNT 54331 a 40998 c 43020 g 54457 t 505 others
ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 193311;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AACTGGATTCGATCAAGTCA 235
|||||
Db 41350 AACTGGATTCGATTCAGTCA 41371

RESULT 7
AC091790

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AC091790 195037 bp DNA linear HTG 12-APR-2002
Sus scrofa clone RP44-278G22, WORKING DRAFT SEQUENCE, 4 ordered
pieces.
AC091790
AC091790.2 GI:20143538
HTG; HTGS_PHASE2; HTGS_DRAFT.
Sus scrofa (pig)
Sus scrofa

REFERENCE
AUTHORS

1 (bases 1 to 195037)
Akheri,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Grantle,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Ido,J.R., Karlins,E., Laric,P.,
Lee-Uin,S.-Q., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskerl,B., Mastrian,S.D.,
McCloskey,J.C., McDowell,J., Pagurigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantirlop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgenev,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 195037)
Green,E.D.
Direct Submission
Submitted (07-JUN-2001) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
3 (bases 1 to 195037)
Green,E.D.
Direct Submission
Submitted (12-APR-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Galtersburg, MD 20877, USA
On Apr 12, 2002 this sequence version replaced gi:14327783.

COMMENT

----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nhgri.nih.gov
----- Project Information
Center project name: cae
Center clone name: 278G22

The sequence data in this record represents an 'enhanced'
version of a Phase 2 submission. Specifically, the indicated
order and orientation of each sequence contig has been
established using one or more of the following: read-pair
data from individual subclones, overlaps with neighboring
clones, alignment with available reference sequence (e.g.,
human), and/or confirmation by PCR testing. In addition,
the sequence assembly is based on at least 8X average
coverage in Q20 bases and has been reviewed to rule out
gross misassemblies, the low-quality ends of sequence
contigs have been trimmed away, and each base is associated
with a Phrap-derived quality score.

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 193528 bases at least Q40
Consensus quality: 194261 bases at least Q30
Consensus quality: 194570 bases at least Q20
Insert size: 144000; agarose-fp
Insert size: 194737; sum-of-contigs
Quality coverage: 11.89x in Q20 bases; sum-of-contigs
Quality coverage: 8.79x in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.

* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
1 14284: contig of 14284 bp in length
* 14285 14384: gap of unknown length
* 14385 26042: contig of 11658 bp in length
* 26043 26142: gap of unknown length
* 26143 64611: contig of 38469 bp in length
* 64612 64712: gap of unknown length
* 64712 195037: contig of 130326 bp in length.
Location/Qualifiers
1. 195037
/organism="Sus scrofa"
/mol_type="genomic DNA"
/db_xref="taxon:9823"
/clone="RP44-278G22"
/clone_11b="RP44"
misc_feature
1. 45776
/note="clone overlaps with GenBank Accession Number
AC092194 clone RP44-309D18 (center project name cad)"
misc_feature
1. 14284
/note="assembly_fragment
clone_end:17
vector_side:left"
misc_feature
14385..26042
/note="assembly_fragment"
misc_feature
26143..64611
/note="assembly_fragment"
misc_feature
64712..195037
/note="assembly_fragment
clone_end:sp6
vector_side:right"
misc_feature
102196..195037
/note="clone overlaps with GenBank Accession Number
AC091791 clone RP44-497D19 (center project name caf)"
BASE COUNT 51409 a 42593 c 44365 g 56370 t 300 others
ORIGIN
Query Match 1.5%; Score 22; DB 2; Length 195037;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 660 GCTTCCAGATGACATGACGAC 681
|||||
DB 109375 GCTTCCAGATGACATGACGAC 109396

RESULT 8
AC140183 205682 bp DNA linear HTG 23-FEB-2003
LOCUS MUS musculus chromosome UNK clone RP24-350N4, WORKING DRAFT
DEFINITION
SEQUENCE, 3 unordered pieces.
AC140183
AC140183.1 GI:28475416
HTG: HTGS_PHASE1, HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
MUS musculus (house mouse)
SOURCE
MUS musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 205682)
McPherson, J.D. and Waterston, R.H.
The sequence of Mus musculus clone
Unpublished
2 (bases 1 to 205682)
McPherson, J.D. and Waterston, R.H.
Direct Submission
Submitted (23-FEB-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>

Contact: submissions@watson.wustl.edu
----- Project Information -----
Center project name: M_B80350N04
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 204725 bases at least Q40
Consensus quality: 204994 bases at least Q30
Consensus quality: 205269 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 38205: contig of 38205 bp in length
* 38206 38305: gap of unknown length
* 38306 99264: contig of 60959 bp in length
* 99265 99365: gap of unknown length
* 99365 205682: contig of 106318 bp in length.
Location/Qualifiers
1. 205682
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="XNK"
/clone="RP24-350N4"
misc_feature
1. 38205
/note="assembly_name:Contig15"
misc_feature
38306..99264
/note="assembly_name:Contig16"
misc_feature
99365..205682
/note="assembly_name:Contig17"
BASE COUNT 64084 a 45447 c 42625 g 53326 t 200 others
ORIGIN
Query Match 1.5%; Score 22; DB 2; Length 205682;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1248 CAAATTTAAGCTAATTCGATT 1269
|||||
DB 67895 CAAATTTAAGCTAATTCGATT 67916

RESULT 9
AC118916 229331 bp DNA linear HTG 19-NOV-2002
LOCUS Rattus norvegicus clone CH230-177J9, WORKING DRAFT SEQUENCE.
DEFINITION
AC118916
AC118916.4 GI:25090461
HTG: HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
KEYWORDS
Rattus norvegicus (Norway rat)
SOURCE
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 229331)
Muzny, D.Marie., Metzker, M.Lee., Abramzon, S., Adams, C., Alder, J.,
Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D.,
Anyalebech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,
Biswal, K., Blais, J., Blankenburg, K., Blyth, P., Brown, M.,
Bryant, N., Buhay, C., Burch, P., Butrell, K., Calderon, E.,
Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,

AUTHORS
 Muzy, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Angiano, D., Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Ceasar, H., Center, E., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J., Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Devila, M., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durdin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Friser, C., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Geregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, N., Guera, W., Gunaratne, P., Healand, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S., L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kwis, C., Kraft, C., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, L., Louised, H., Lozano, R., J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhley, S., McLeod, M., McNeill, T., Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Morgan, M., Morris, K., Morris, S., Mundasa, M., Murphy, M., Naier, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwokelam, O., Okunolu, G., Olariunsaagoo, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankoch, C., Plummer, F., Polidexter, A., Popovic, D., Primus, E., Pu, L., L., Puzo, M., Quiroz, J., Rachin, E., Reeves, K., Regier, M., A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S., J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Slisson, I., Sitter, C., D., Smajs, D., Sneed, A., Sodergren, E., Song, X., Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Taber, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczyk, R., Woodson, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, S., Zhao, S., Dunn, O., von Niederhausern, A., Weiss, R., Smith, D., R., Holt, R., A., Smith, H., O., Weinstein, G., and Gibbs, R. A.

TITLE
 Direct Submission
 Unpublished
 2 (bases 1 to 231442)

REFERENCE
 Submitted (15-SEP-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 231442)
 Rat Genome Sequencing Consortium.

AUTHORS
 Submitted (09-MAY-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT
 On May 9, 2003 this sequence version replaced g1:22771413. The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'config-scaffold'). Within each config-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a config-scaffold that consist entirely of whole

genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
 Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GBMC
 Center clone name: CH230-6121
 ----- Summary Statistics
 Assembly program: Atlas;
 Consensus quality: 207859 bases at least Q40
 Consensus quality: 211922 bases at least Q30
 Consensus quality: 214681 bases at least Q20
 Estimated insert size: 220165; sum-of-contigs estimation
 Quality coverage: 7x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 1 contigs. Gaps between the contigs
 * are represented as runs of N. The order of the pieces
 * is believed to be correct as given, however the sizes
 * of the gaps between them are based on estimates that have
 * provided by the submitter.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 1 231442: contig of 231442 bp in length.
 Location/Qualifiers
 1..231442
 /organism="Rattus norvegicus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10116"
 /clone="CH230-6121"
 1..1860
 /note="wgs_end_extension
 clone_end:17"
 588..6759
 /note="clone_boundary
 clone_end:17
 site:EcoRI
 end_sequence:BH360141"
 229224..229921
 /note="clone_boundary
 clone_end:Sp6
 site:EcoRI
 end_sequence:BH360143"
 BASE COUNT 63051 a 48424 c 46447 g 58733 t 14787 others
 ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 231442;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

misc-feature
 misc-feature
 misc-feature

Qy 214 AACTGCGATTCGATTCAGTCA 235
 ||||||||||||||||||||
 Db 72367 AACTGCGATTCGATTCAGTCA 72346

RESULT 11
 AC128235 253961 bp DNA linear HTG 21-SEP-2002
 LOCUS AC128235/c
 DEFINITION Rattus norvegicus clone CH230-38615, *** SEQUENCING IN PROGRESS
 *** 4 unordered pieces.
 AC128235
 AC128235.2 GI:23265191
 VERSION HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
 KEYWORDS Rattus norvegicus (Norway rat)
 SOURCE Rattus norvegicus
 ORGANISM Rattus norvegicus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
Rattus.
1 (bases 1 to 253961)

Muzny,D,Marie, Metzger,M, Lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsprouks,S, Amin,A, Anguiano,D, Anyalebech,V, Aoyagi,A, Ayodeji,M, Baca,E, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Bisvalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,E, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,D, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,H, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,T, Fraser,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,M, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamill,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hulak,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolyet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Kovar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawlinne,S, McLeod,M,P, McNeill,T,Z, Meenen,E, Milosavljevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morán,M, Morris,K, Morris,S, Munidas,M, Murphy,M, Natl,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwaekelemel,O, Okunnu,G, Olarnpusgon,A, Pal,S, Parks,K, Pasternak,S, Paul,H, Perez,A, Perez,L, Pfankoch,C, Plopper,F, Poindexter,A, Popovic,D, Primus,E, Pu,L, Puazo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reilgh,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,J, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatsman,S, Shen,H, Shetty,J, Shvartsbeyn,A, Sisson,I, Sitter,C,D, Smajs,D, Sneed,A, Sodergren,E, Song,X,-Z, Sorelle,R, Sosa,J, Steidle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Usmani,K, Valas,R, Vera,V, Villasana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,F, Williams,G, Willson,R, Wleczyk,R, Wooden,H, Worley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, Von Niederhausern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
REFERENCE
AUTHORS
TITLE
JOURNAL

Unpublished
2 (bases 1 to 253961)
Worley,K.C.
Direct Submission
Submitted (19-Jun-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 253961)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (21-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Sep 21, 2002 this sequence version replaced gi:21908845.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). As a result, the sequence may extend beyond the ends of the clone and there may be contigs that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only

contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: GYBV

Center clone name: CH230-38G15

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 214169 bases at least Q40

Consensus quality: 218921 bases at least Q30

Consensus quality: 222222 bases at least Q20

Estimated insert size: 237536; sum-of-contigs estimation

Quality coverage: 4x in Q20 bases; sum-of-contigs estimation

----- NOTE: Estimated insert size may differ from sequence length

(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).

* NOTE: This is a 'working draft' sequence. It currently

* consists of 4 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

* 1 249583: contig of 249583 bp in length

* 249584 249683: gap of unknown length

* 249684 250819: contig of 1136 bp in length

* 250820 250919: gap of unknown length

* 250920 252666: contig of 1747 bp in length

* 252667 252766: gap of unknown length

* 252767 253961: contig of 1195 bp in length.

FEATURES

source 1.253961

Location/Qualifiers

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-38G15"

1.1470

/note="wgs_end-extension"

clone_end:77

4785..5385

/note="clone boundary"

clone_end:77

site:ECORI

end.sequence:BH277815"

174876..175881

/note="wgs_contig"

177111..178416

/note="wgs_contig"

63786 a 51180 c 48941 g 61233 t 28821 others

BASE COUNT

ORIGIN

Query Match 1.5%; Score 22; DB 2; Length 253961;

Best Local Similarity 100.0%; Pred. No. 3.7;

Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 214 AACTGGATTGATTCAGTCA 235

DB 217452 AACTGGATTGATTCAGTCA 217431

RESULT 12

BD146582

LOCUS

BD146582 872 bp DNA linear

DEFINITION

Primer for synthesizing full-length cDNA and use thereof.

ACCESSION

BD146582.1 GI:27852340

KEYWORDS

UP 2002191363-A/1425.

SOURCE

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 872)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 1425 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/1425
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
 PI KEIICHI NAGAI,TETSUJI OTSUKI
 PC
 C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
 10,
 C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
 PC
 Primer for synthesizing full-length cDNA and use thereof FH key
 Location/Qualifiers
 FT source 1..872
 1..872 /organism="Homo sapiens (human)".
 Location/Qualifiers
 FT source 1..872
 1..872 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 285 a 138 c 177 g 263 t 9 others
 ORIGIN
 Query Match 1.4%; Score 21; DB 6; Length 872;
 Best Local Similarity 100.0%; Pred. No. 8.3;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 873 TGATCAATTAGCAAAAGATCC 893
 DB 747 TGATCAATTAGCAAAAGATCC 767
 RESULT 13
 BDI59393
 LOCUS BDI59393 2202 bp DNA linear PAT 17-JAN-2003
 DEFINITION Primer for synthesizing full-length cDNA and use thereof.
 ACCESSION BDI59393.1 GI:27865151
 VERSION BDI59393.1
 KEYWORDS JP 2002191363-A/14236.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2202)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 14236 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/14236
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
 PI KEIICHI NAGAI,TETSUJI OTSUKI
 PC
 C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
 10,
 C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
 PC
 Primer for synthesizing full-length cDNA and use thereof FH key
 Location/Qualifiers
 FT source 1..2202

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 872)
 Ota,T., Isogai,T., Nishikawa,T., Hayashi,K., Saito,K., Yamamoto,J., Ishii,S., Sugiyama,T., Wakamatsu,A., Nagai,K. and Otsuki,T.
 Primer for synthesizing full-length cDNA and use thereof
 Patent: JP 2002191363-A 1425 09-JUL-2002;
 HELIX RESEARCH INSTITUTE
 OS Homo sapiens (human)
 PN JP 2002191363-A/1425
 PD 09-JUL-2002
 PF 28-JUL-2000 JP 2000280990
 PI TOSHIO OTA,TAKAO ISOGAI,TETSUO NISHIKAWA,KOJI HAYASHI,KAORU
 PI SAITO,
 PI JUNICHI YAMAMOTO,SHIZUKO ISHII,TOMOYASU SUGIYAMA,AI WAKAMATSU,
 PI KEIICHI NAGAI,TETSUJI OTSUKI
 PC
 C12N15/09,C07K14/47,C07K16/18,C12N1/15,C12N1/19,C12N1/21,C12N5/PC
 10,
 C12P21/02,C12Q1/68//C12P21/08,G06F17/30,C12N15/00,C12N5/00 CC
 PC
 Primer for synthesizing full-length cDNA and use thereof FH key
 Location/Qualifiers
 FT source 1..872
 1..872 /organism="Homo sapiens (human)".
 Location/Qualifiers
 FT source 1..872
 1..872 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 BASE COUNT 285 a 138 c 177 g 263 t 9 others
 ORIGIN
 Query Match 1.4%; Score 21; DB 6; Length 2202;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 873 TGATCAATTAGCAAAAGATCC 893
 DB 746 TGATCAATTAGCAAAAGATCC 766
 RESULT 14
 AK022163
 LOCUS AK022163 2202 bp mRNA linear PRI 01-AUG-2002
 DEFINITION Homo sapiens cDNA FLJ12101 fls, clone HEMDB1002683.
 ACCESSION AK022163
 VERSION AK022163.1 GI:10433497
 KEYWORDS cDNA cloning; full length insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 2202)
 Isogai,T., Ota,T., Hayashi,K., Sugiyama,T., Otsuki,T., Suzuki,Y., Nishikawa,T., Nagai,K., Sugano,S., Takahashi-Fujii,A., Hara,H., Tanase,T., Nomura,Y., Togiya,S., Komai,F., Hara,R., Takeuchi,K., Arita,M., Nebekura,T., Ishii,S., Kawai,Y., Saito,K., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagahara,K., Masuno,Y. and Oshima,A.
 NEDO human cDNA sequencing project
 Unpublished
 2 (bases 1 to 2202)
 Isogai,T. and Otsuki,T.
 Direct Submission
 Submitted (23-AUG-2000) Takao Isogai, Helix Research Institute, Genomics Laboratory; 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:genomics@hli.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
 NEDO human cDNA sequencing project supported by Ministry of International Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology; cDNA library construction; 5'- & 3'-end one pass sequencing and clone selection; Helix Research Institute (supported by Japan Key Technology Center etc.) and Department of Virology, Institute of Medical Science, University of Tokyo.
 FEATURES
 source 1..2202
 1..2202 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="HEMDB1002683"
 /issue_type="whole embryo, mainly body"
 /clone_lib="HEMB1"
 /dev_stage="embryo, 10 weeks"
 /note="cloning vector: pME18SFL3"
 BASE COUNT 703 a 416 c 488 g 595 t
 ORIGIN
 Query Match 1.4%; Score 21; DB 9; Length 2202;
 Best Local Similarity 100.0%; Pred. No. 9;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 873 TGATCAATTAGCAAAAGATCC 893
 DB 746 TGATCAATTAGCAAAAGATCC 766
 RESULT 15

DMC137E7/c 11922 bp DNA linear INV 27-APR-1999
 LOCUS DMC137E7
 DEFINITION Drosophila melanogaster cosmid clone 137E7.
 ACCESSION AL021108
 VERSION AL021108.1 GI:4164288
 KEYWORDS
 SOURCE Drosophila melanogaster (fruit fly)
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 11922)
 Ferraz,C., Vidal,S., Brun,C., Bucheton,A. and Demaille,J.G.
 Sequencing the distal X chromosome of Drosophila melanogaster
 Unpublished
 Institut de Genetique Humaine -UPR 1142- CNRS, 141,rue de la
 Cardonille, 34396 Montpellier Cedex 5, France
 2 (bases 1 to 11922)
 Benos,P.
 Direct Submission
 Submitted (26-APR-1999) European Drosophila Genome Sequencing
 Consortium
 On Jan 19, 1999 this sequence version replaced gi:3947671.
 Sequence submitted by Takis Benos, EMBL Outstation - The EBI,
 Hinxton, Cambridge, CB10 1SD, U.K.
 E-mail: benoseebi.ac.uk on behalf of the European Drosophila Genome
 Sequencing Consortium. For further information see the European
 Drosophila Genome Sequencing Consortium's web site:
 http://edgp.ebi.ac.uk/.
 Coding sequences are predicted from computer analysis, using both
 gene and CDS prediction programs and matches to other sequences.
 These predictions and matches have been evaluated by the annotators
 and may have been refined by hand (in which case a GeneFinder
 prediction will have no score. The annotators have also used their
 judgement on what matches to represent in this record. A far more
 complete annotation record is available from FlyBase
 (http://flybase.bio.indiana.edu/) through the FlyBase Annotation
 Object linked by the db_xref qualifier in the Feature Table.
 IMPORTANT: This sequence is NOT necessarily the entire insert of
 clone 137E7. It may be shorter, since we are minimising the
 overlap between clones to 100 bases, by trimming them. Clone 137E7
 overlaps to the left with clone 30B7
 The true left end of clone 137E7 is at position 4266 of clone 30B7
 Clone 137E7 overlaps to the right with clone 131F2 Sequence in
 absolute orientation with respect to chromosome The syntax for the
 representation of annotation used in this record is documented at:
 ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation.README
 Clone=137E7; Contig ID=1; Length=11922; Status=Finished.
 Location/Qualifiers
 1..11922
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /clone="cosmid 137E7"
 4333..4524
 /gene="EG:137E7.1"
 4333..4524
 /gene="EG:137E7.1"
 /note="EG:137E7.1"
 /note="EG:137E7.1"
 /prediction="method:'genefinder',
 version:'084', score:'12.36'",
 /match="desc:'LD19625.Sprime LD Drosophila melanogaster
 embryo Bluescript Drosophila melanogaster cDNA clone
 LD19625.Sprime. mRNA sequence'", species:'Drosophila
 melanogaster (fruit fly)', ranges:(query:3905..4079,
 target:EMBL:AA540213:1..175, score:'848.00')",
 method:'blastn', version:'1.4.9')"
 the EST EMBL:AA540213 comes from the 5' UTR"
 /codon_start=1
 /protein_id="CA15947.1"
 /db_xref="GI:2760070"
 /db_xref="FLYBASE:FBgn0023543"
 /db_xref="SPTRMBL:O46049"

/translation="MEQITNYFEHNDYVSFSVALPCVVIAYGFLQVASGLACKSK
 SSMPCVCFHLTNOKRIAD"
 BASE COUNT 3384 a 2429 c 2431 g 3678 t
 ORIGIN
 Query Match 1.4%; Score 21; DB 3; Length 11922;
 Best Local Similarity 100.0%; Pred.No. 11;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 409 AACGAAGATACCAAAATGAT 429
 ||||||||||||||||||||
 DB 9602 AACGAAGATACCAAAATGAT 9582
 Search completed: September 2, 2003, 19:33:31
 Job time : 5691 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: September 2, 2003, 20:31:10 : Search time 3066 seconds

(without alignments)
3947.694 Million cell updates/sec

Title: US-09-872-523-5
Perfect score: 498
Sequence: 1 MSEFLKIVRANKSDRKLDK.....PIGWCEANSYLVQPPKKNY 498

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 22781392 seqs, 12152238056 residues

Word size: 1

Total number of hits satisfying chosen parameters: 45562604

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters: -DEV=x1h
-MODER=frame.p2n.model -DEV=x1h
-Q/cgn2.1/USPTO.spool/US09872523/runat.02092003.073002.27259/app.query.fasta.1.647
-DB-EST -QFMT=fastap -SUFFIX=ol1.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc
-NORM-ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09872523 @CGN1.1.2810 @runat.02092003.073002.27259 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEDUERY -NEG.SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMECUT=120 -WRN_TIMECUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estinu:*
5: em_estlov:*
6: em_estpl:*
7: em_estro:*
8: em_hlc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hlc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_esttom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pin:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	210	42.2	653	9	AU204939
2	186	37.3	581	12	BJ101701
3	164	32.9	557	12	BJ124772
4	148	29.7	519	12	BJ115564
5	138	27.7	768	12	BJ143186
6	135	27.1	761	12	BJ129361
7	128	25.7	607	14	CB404801
8	119	23.9	360	14	C42305
9	112	22.5	380	9	AV197399
10	111	22.3	677	9	AU216908
11	108	21.7	686	12	BJ153087
12	102	20.5	340	9	AU111588
13	72	14.5	300	9	AU116296
14	70	14.1	300	9	AU115428
15	62	12.4	337	9	AU112450
16	22	4.4	505	14	CB385690
17	10	2.0	504	28	AO605493
18	9	1.8	196	29	AG241054
19	9	1.8	391	29	AG229361
20	9	1.8	456	12	BI097227
21	9	1.8	494	12	BI097088
22	9	1.8	502	10	BG668175
23	9	1.8	534	28	AO415204
24	9	1.8	608	12	BI097296
25	9	1.8	610	12	BI097211
26	9	1.8	626	12	BI097047
27	9	1.8	637	12	BI097165
28	9	1.8	644	12	BI097081
29	9	1.8	655	28	AO389984
30	9	1.8	663	12	AO389894
31	9	1.8	693	12	BI096384
32	9	1.8	701	28	BI097080
33	9	1.8	704	12	BI096997
34	9	1.8	715	29	AG071711
35	9	1.8	722	14	CB445305
36	9	1.8	738	14	CB858090
37	9	1.8	739	14	CB445578
38	9	1.8	793	12	BI096871
39	9	1.8	877	14	CD303479
40	9	1.8	917	14	CD359252
41	8	1.6	93	9	A1206304
42	8	1.6	184	28	B64089
43	8	1.6	188	9	A1559920
44	8	1.6	190	9	AA593206
45	8	1.6	190	28	AO081464

ALIGNMENTS

RESULT 1
LOCUS AU204939 653 bp mRNA linear EST 17-JUL-2001
DEFINITION AU204939 unpublished oligo-capped cDNA library, stage I4
ACCESSION AU204939
VERSION AU204939.1 GI:14836792
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea
; Rhabditidae; Peioderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 653)

AUTHORS Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.

TITLE A complementary view of the C.elegans genome

JOURNAL Unpublished

COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
Location/Qualifiers
1. 653
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK843e05"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L4"
/clone_lib="unpublished oligo-capped cdna library, stage L4"

BASE COUNT 216 a 116 c 136 g 185 t

ORIGIN

Alignment Scores:
Pred. No.: 4.79e-202 Length: 653
Score: 210.00 Matches: 210
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 42.17% Indels: 0
DB: Gaps: 0

US-09-872-523-5 (1-498) x AU204939 (1-653)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
DB 22 ATGCTGAATTTCTGAAAATTGTCAGACTAACAAAATCGGACAGAAAACCTCGATTAG 81
QY 21 ThrTYrLeuTPGluSerTYrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB 82 ACCTACTGTGGGAATCCTATTACATCAGTTCCGAAAGAAAACCTTTCTTCTTCCA 141
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluLysI 60
DB 142 GTTGACGATTCACATCGTAACCTTACCTTAATTTAAACGATGCGTAAGACGAGACTT 201
QY 61 IlePheGluThrValValHisAspTYrAspLysAsnCysAspSerIleGlnValArgTrp 80
DB 202 ATCTTCGAACAGTGGCTCATGATTATGACAGAACTGCGATTCCATTCAGTCAGATGG 261
QY 81 PheAlaArgIleGluLysValCysGlyTYrArgValLeuAlaGlnPheIleGlyAlaAsp 100
DB 262 TTTGACGATTCGAAAAGTTTGGCGATACAGAGTTTGGCTCATGATTATCGAGCTGAC 321
QY 101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaMet 120
DB 322 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTGGCAACGCCCAATG 381
QY 121 SerAspProAsnMetAspLysIleValTYrAlaProProLeuAlaIleAsnGluLys 140
DB 382 AGTGATCCCAATATGATTAATGTTATATGCTCCGCCGCTTGCAATCAACAAGAAATTC 441
QY 141 GlnAsnAspMetValAsnTYrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
DB 442 CAATATGATATGTAATTAATTAATTAATTCATTCATTCATTCATTCATTCATTCATTC 501
QY 161 SerLeuSerProLysPheAspGluLysAlaLeuLeuSerLysHisArgPheLysVal 180
DB 502 TCGTGCTCCCAAAATTCGATGAAGGAGCTCTCTAAGCAACATCGTTTCAAAAGTT 561
QY 181 GlyAlaArgLeuGluLeuLeuAsnTYrSerAsnSerThrGluIleArgValAlaArgTle 200

|||||

DB 562 GGACACACTCTTGAACATTTAATTTACTCCATTTCTACTGAATAACCGTACGCGAATT 621

QY 201 GlnGluIleCysGlyArgArgMetAsnVal 210
|||||

DB 622 CAAGAATATGCTGGACGACGAATGATGTA 651

RESULT 2
BU101701 581 bp mRNA linear EST 18-JAN-2002
LOCUS BU101701 unpublished oligo-capped cdna library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cdna clone YK1029h03 5', mRNA sequence.
ACCESSION BU101701 GI:18244371
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 581)
REFERENCE Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y. and Sugano, S.
A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.

FEATURES
Location/Qualifiers
1. 581
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="YK1029h03"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cdna library, C. elegans L1 stage"

BASE COUNT 169 a 103 c 123 g 166 t

ORIGIN

Alignment Scores:
Pred. No.: 8.65e-178 Length: 581
Score: 186.00 Matches: 186
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 37.35% Indels: 0
DB: Gaps: 0

US-09-872-523-5 (1-498) x BU101701 (1-581)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
DB 23 ATGCTGAATTTCTGAAAATTGTCAGACTAACAAAATCGGACAGAAAACCTCGATTAG 82
QY 21 ThrTYrLeuTPGluSerTYrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
DB 83 ACCTACTGTGGGAATCCTATTACATCAGTTCCGAAAGAAAACCTTTCTTCTTCCA 142
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluLysI 60
DB 143 GTTGACGATTCACATCGTAACCTTATTTAAGAAATGCGTGAAGGAAGACTT 202
QY 61 IlePheGluThrValValHisAspTYrAspLysAsnCysAspSerIleGlnValArgTrp 80
DB 203 ATCTTCGAACAGTGGCTCATGATTATGACAGAACTGCGATTCCATTCAGATCG 262
QY 81 PheAlaArgIleGluLysValCysGlyTYrArgValLeuAlaGlnPheIleGlyAlaAsp 100

Db 263 TTGGCAGAAATGAAAAAGTTGGCGATACAGAGTCTCGCTCAGTTATCGAGCTGAC 322
 QY 101 ThrlypPheTyrLeuAsn1lleuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
 Db 323 ACGAAATTTTGGCTCAATATTTTATTCGACGATATGTTGGTGGCAAAAGCCCGCAATG 382
 QY 121 SerAspProAsnMetAspLysIleVal1TyrAlaProLeuAlaIleAsnGluGlyTyr 140
 Db 383 AATGATCCCAATATGATGATAAATGTATGTCTCCGCCCTTGCAATCAACGAGAAATAC 442
 QY 141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
 Db 443 CAAATGATATGATGTAATATGTAATATGATGATGCGCAATCGTCGCCCAACT 502
 QY 161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
 Db 503 TCGCTGTCTCCAAATTCGATGAAAGGAGGCTCTCCTTAAGCAAGCATGTTTCAAAAGTT 562
 QY 181 GlyGlnArgLeuGluLeu 186
 Db 563 GGACAAACGTCTTGAACTA 580

RESULT 3
 BJI24772 557 bp mRNA linear EST 30-MAY-2003
 LOCUS BJI24772 unpublished oligo-capped cDNA library, C. elegans L1 stage
 DEFINITION Caenorhabditis elegans cDNA clone yk1315b11 5', mRNA sequence.
 ACCESSION BJI24772 GI:31246008
 VERSION BJI24772.2 GI:31246008
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidae
 ; Rhabditidae; Peloderinae; Caenorhabditis.
 REFERENCE 1 (bases 1 to 557)
 Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
 and Sugano, S.
 A complementary view of the C. elegans genome
 Unpublished
 On Jan 23, 2002 this sequence version replaced gi:18284913.
 TITLE JOURNAL
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers

FEATURES
 source 1..557
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK1315B11"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 180 a 180 g 158 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1 61e-155 Length: 557
 Score: 164.00 Matches: 177
 Percent Similarity: 99.448 Conservative: 0
 Best Local Similarity: 99.448 Mismatches: 0
 Query Match: 32.938 Indels: 1
 DB: 12 Gaps: 0

US-09-872-523-5 (1-498) x BJI24772 (1-557)

QY 2 SerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLysThr 21
 Db 26 TCTGATTTCTGAAATTTCTCAGAGCTAACCAAAAATCGACAGAAACATCGATAGAC 85
 QY 22 TyrLeuTyrGluSerTyrLeuHisGlnPheGluLysGly-LysThrSerPheIleProVa 41
 Db 86 TACTTGTGGGAATCCATATTTACATCATGTTTCGAGAAAAGNAAAAAACTTTTCATCTCACT 145
 QY 41 IGLuAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyAla11 61
 Db 146 TGAAGCATTCATTCGTAACCTTACACTTAATTTTAAAGAAATGCGTGAAGGAAGGACTTAT 205
 QY 61 ePheGluThrValAlaHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyrP 81
 Db 206 CTTTCAAAACAGTGTCCATGATTAATACAAAGAACTCCATTCGATTCAGTCAGATGATT 265
 QY 81 eAlaArgIleGluLysValLysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAspTh 101
 Db 266 TGCACGAATTTGAAAAAGTTTGGGATACAGAGTTCGGCTCAGTTTATCGAGCTGACAC 325
 QY 101 rlyspPheTyrPleuAsn1lleuSerAspAspMetPheGlyLeuAlaAsnAlaIleMetSe 121
 Db 326 GAAATTTTGGCTCAATATTTTATTCGACGATATGTTGGTTGGCAAAAGCCGCAATGAG 385
 QY 121 rAspProAsnMetAspLysIleVal1TyrAlaProProLeuAlaIleAsnGluGlyTyrG1 141
 Db 386 TGAATCCCAATATGCAATAAATGTATATCTCCGCCCTTGCAATCAACGAGAAATACCA 445
 QY 141 nAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThrSe 161
 Db 446 AATGATATGATGTAATATGTAATATGATGATGATGATGATGATGATGATGATGATGAT 505
 QY 161 rLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPhe 178
 Db 506 GCTGTCTCCAAATTCGATGAAAGGAGGCTCTCCTTAAGCAAGCATGCTTTC 557

RESULT 4
 BJI15564 519 bp mRNA linear EST 23-JAN-2002
 LOCUS BJI15564 unpublished oligo-capped cDNA library, C. elegans L1 stage
 DEFINITION Caenorhabditis elegans cDNA clone yk1192n09 5', mRNA sequence.
 ACCESSION BJI15564 GI:18275658
 VERSION BJI15564.1 GI:18275658
 KEYWORDS EST.
 SOURCE Caenorhabditis elegans
 ORGANISM Caenorhabditis elegans
 Eukaryota; Metazoa; Nematoda; Chromodorea; Rhabditida; Rhabditoidae
 ; Rhabditidae; Peloderinae; Caenorhabditis.

REFERENCE 1 (bases 1 to 519)
 Kohara, Y., Shin-I, T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.
 and Sugano, S.
 A complementary view of the C. elegans genome
 Unpublished

TITLE JOURNAL
 COMMENT Contact: Tadasu Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.
 Location/Qualifiers

FEATURES
 source 1..519
 /organism="Caenorhabditis elegans"
 /mol_type="mRNA"
 /strain="N2"
 /db_xref="taxon:6239"
 /clone="YK1192N09"
 /sex="hermaphrodite"
 /tissue_type="whole animal"
 /dev_stage="L1"
 /clone_lib="unpublished oligo-capped cDNA library, C.
 elegans L1 stage"
 BASE COUNT 170 a 91 c 109 g 149 t

US-09-872-523-5 (1-498) x BJI15564 (1-519)

ORIGIN

Alignment Scores:

Alignment Scores: 2.4e-139 Length: 519
Pred. No.: 148.00 Matches: 148
Score: 148.00 Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 29.72% Gaps: 0

DB:

US-09-872-523-5 (1-498) x BJ115964 (1-519)

OY 1 Metserglupheleuylslllevalargalalsnlyssersaparglyleuaspys 20
DB 16 ATGCTCGAATTCGAAAATGTCAGACCTACAAAATGCGACAGAACTCATAG 75
OY 21 Thrlyleutrgluserlyrleuhsiglnpheglulysglylsthserphelepro 40
DB 76 ACCTACTGTGGGAATCCTATTACATCAGTTCGAGAAAGAAAACCTCTTCATTCCA 135
OY 41 Valglualapheasnargasnleuthrvalasnpheasnnglucysvallysgluylval 60
DB 136 GTTGACGACATTCACATCGTAACCTTACAGTATTTTAAAGATCGGTGAAGAGAGTT 195
OY 61 Ilephegluphrvalvalhisaprtyrasplysasncysaspserrileglvalargtrp 80
DB 196 ATCTTCGAACAGAGTGTCATGATTAAGACAAAGACTCGATTCGATTCAGTCAGATGG 255
OY 81 Phealaargileglulysvalcysgllytrargvalleualaglnpheileglvalaasp 100
DB 256 TTTCGACGAAATGAAAAGTTTGGCGATACAGAGTTCGCTCATGTTATCGAGCTGAC 315
OY 101 Thrlyspheirpleuasnilleuuseraspasphetpheglyleuualasnlaalamet 120
DB 316 ACGAAATTTTGGCCCAATATTTTATCGGACGATATGTTGGTTGGCAAGCCCGCAATG 375
OY 121 Seraspproasnmetaspyslllevaltyralaproproleualailasnglulutyr 140
DB 376 AGTATGCCAATAGGATTAATAATGTTATGCTCCGCCCTTGCAATCAAGAAATATC 435
OY 141 Glinsnaspmetvalasntrval 148
DB 436 CAAATGATGATGTAATATGTA 459

RESULT 5
BJ143186 768 bp mRNA linear EST 23-JAN-2002
LOCUS BJ143186 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1192h09 3', mRNA sequence.
ACCESSION BJ143186
VERSION BJ143186.1 GI:18303352
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
1 (bases 1 to 768)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES
SOURCE
1..768
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"

/db_xref="taxon:6239"

/clone="yk1192h09"

/sex="hermaphrodite"

/tissue_type="whole animal"

/dev_stage="L1"

/clone_lib="unpublished oligo-capped cDNA library, C.

elegans L1 stage"

BASE COUNT

232 a 151 c 146 g 237 t 2 others

ORIGIN

US-09-872-523-5 (1-498) x BJ143186 (1-768)

OY 283 Argtyrpsersaspvalthrphespolnleualalysaspproilleasppromet 302
DB 768 AGATATGACTCGACGACGATCATTTGATCAATTAGCAAAAGATCCATTTGATCCCATG 709
OY 303 Iletrparglusvallyvaliglylnlysphegluleuileaspproleualagln 322
DB 708 ATTTGGAAAGAGTTAAGTTGACCAAAAGTTTGAACCTCATCGACCCCTTGCTCAGCA 649
OY 323 Pheasnasnleuhsivalalaserlleuulysphecyslystrnglulgllyrleuile 342
DB 648 TTCAATTAACCTCCACGCTGCTTCATCTCAAAATTTTCAAAACGGAAGATATCTTAT 589
OY 343 Valglymetaspglpyproaspalaleuengluspserrpheprollehisilasnasthr 362
DB 588 GTGGAAAGATGAGTGCCAGATGCACCTTGAAACAGTTTCTTATTCATATCAATTAACA 529
OY 363 Phehetpheprovaliglytyralagluptyrasnleuengluleualpyproaspolu 382
DB 528 TTTATGTTCCCAAGTTGTTATGCGGAAAGATTAATTTGGAACCTGTTCCCGCAGATGAG 469
OY 383 Pheylsglythrpheargtrpaspoluptyrleuenglulysgluseralagluthrleupro 402
DB 468 TTCAAGGAAACATTCAGATGAGTGAATATCTTGGAAGAAATCTGCAGAAACCTTACCG 409
OY 403 Leuaspheupheylspprometpserglngluargleuaspyspheylsval 420
DB 408 CTGACCTGTTCAAGCCATGCTCTCCCAAGAGAGATTAGCAAAATTTTAAGTA 355

RESULT 6
BJ129361 761 bp mRNA linear EST 23-JAN-2002
LOCUS BJ129361 unpublished oligo-capped cDNA library, C. elegans L1 stage
DEFINITION Caenorhabditis elegans cDNA clone yk1029h03 3', mRNA sequence.
ACCESSION BJ129361
VERSION BJ129361.1 GI:18289518
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
1 (bases 1 to 761)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
TITLE A complementary view of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Tadasu Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshini@genes.nig.ac.jp.
Location/Qualifiers
FEATURES

```
source
1. .761
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk1029h03"
/sex="hermaphrodite"
/tissue_type="whole animal"
/dev_stage="L1"
/clone_lib="unpublished oligo-capped cDNA library, C. elegans L1 stage"
BASE COUNT 233 a 150 c 144 g 232 t 2 others
ORIGIN

Alignment Scores:
Pred. No.: 5 64e-126 Length: 761
Score: 135.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.11% Indels: 0
DB: 12 Gaps: 0

US-09-872-523-5 (1-498) x BJ129361 (1-761)
QY 286 SerAspAspValThrPheAspGlnLeuAlaLysAspProIleAspPromeIleTyrParg 305
Db 759 TCAGACGACGTCACATTGATTCATTAAGCAAAAGATTCATTCATCCATGATTTGGAGA 700
QY 306 LysValLysValGlnLysPheGlnLeuIleAspProLeuAlaGlnGlnPheAsnAsn 325
Db 699 AAGATTAAAGGTTGACAAAGATTGAGCTCATGACCCNTGGCTCAGCAATTCAAATAC 640
QY 326 LeuHisValAlaSerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMet 345
Db 639 CTCACAGTCGCTCGATTCCTCAATTTTGGCAAAACGACGATCTTATTGGGCAATG 580
QY 346 AspGlyProAspAlaLeuGlnAspSerPheProIleHisIleAsnAsnThrPheMetPhe 365
Db 579 GATGGTCACAGATCAGATTGAGACAGATTTCCTATTCATCATATACATTTAAGTTTC 520
QY 366 ProValGlyTyrAlaGlnLysTyrAsnLeuGlnLeuValProProAspGlnPheLysGly 385
Db 519 CCGATTGGTTATCGGAAAGATTAATTTGGAACTGTTCCGACAGTGAATTCACAAAGA 460
QY 386 ThrPheArgTyrAspGlnLysGlnLeuGlnSerAlaGlnThrLeuProLeuAspLeu 405
Db 459 ACATTGACATGGATGAAATCTTGAGAAAGAAATCTGACAGAAACCTACCGCTTGACTTG 400
QY 406 PheLysPromeIleProSerGlnGlnLysAlaGlnLysPheLysVal 420
Db 399 TTCACGCAATGCCCTCCCAAGAGATTAGACAAATTTAAGGTA 355

RESULT 7
CB404801/ 607 bp mRNA linear EST 15-MAY-2003
LOCUS OSTR02803_1 AD-wrmCDNA Caenorhabditis elegans cDNA, mRNA sequence.
DEFINITION CB404801
ACCESSION CB404801
VERSION 1
KEYWORDS GI:30746528
SOURCE EST.
ORGANISM Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitoidea
; Rhabdilitidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 607)
Reboul,J., Vagillo,P., Rual,J.F., Lamesch,P., Martinez,M., Armstrong
,C.M., Li,S., Jacotot,L., Berth,N., Janky,R., Moore,T., Hudson
,J.R., Hartley,J.L., Brasch,M.A., Vandenhaute,J., Boulton,S.,
Endress,G.A., Jenna,S., Chevett,E., Papasotiriopoulos,V., Tollas,P.P.,
Placek,J., Snyder,M., Huang,R., Chance,M.R., Lee,H.,
Doucette-Stamm,L., Hill,D.E. and Vidal,M.
C. elegans ORFome version 1.1: experimental verification of the
genome annotation and resource for proteome-scale protein
expression
```

```
JOURNAL Nat. Genet., (2003) In press
COMMENT Contact: Vidal M
Marc Vidal Laboratory
Dana Farber Cancer Institute
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
Tel: 617 632 5180
Fax: 617 632 5739
Email: Marc.Vidal@dfci.harvard.edu
Sequence tag of Gateway entry clones. The primers used were
designed on the predicted protein encoding ORF. C. elegans ORFome
cloning project : Contact david.hill@dfci.harvard.edu or
marc.vidal@dfci.harvard.edu
POLYA-No.

FEATURES
source
1. .607
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/sex="hermaphrodite and male"
/tissue_type="whole animal"
/dev_stage="mixed stage"
/clone_lib="AD-wrmCDNA"
/note="The AD-wrmCDNA library was generated with poly(A)+
RNA isolated from both hermaphrodite and male N2 worms of
all larval stages, embryos, adults and dauers and the
subsequent generation of cDNAs by poly(A) priming. The
cDNAs were cloned into pCR86"
BASE COUNT 175 a 131 c 113 g 188 t
ORIGIN

Alignment Scores:
Pred. No.: 5 33e-119 Length: 607
Score: 128.00 Matches: 128
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 25.70% Indels: 0
DB: 14 Gaps: 0

US-09-872-523-5 (1-498) x CB404801 (1-607)
QY 293 GlnLeuAlaLysAspProIleAspPromeIleTyrArgLysValLysGlnLys 312
Db 573 CAATTAGCAAAAGATTCATTCATCCATGATTTGGAGAAAGTTAAGTTGACAAAG 514
QY 313 PheGlnLeuIleAspProLeuAlaGlnGlnPheAsnAsnLeuHisValAlaSerIleLeu 332
Db 513 TTGGAGCTCATGACCCCTGGCTGACCAATTCATTAACCTCCACGCTGCTGATCTC 454
QY 333 LysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProAspAlaLeuGln 352
Db 453 AAATTTTGCAAAAGTGAAGATTCATTATGTGGAAATGATGTCAGATGCACTTGAA 394
QY 353 AspSerPheProIleHisIleAsnAsnThrPheMetPheProValGlyTyrAlaGlnLys 372
Db 393 GACAGTTTCTTATTCATATCATTAATATATATATGTTCCAGTTGGTTATGCGAAAG 334
QY 373 TyrAsnLeuGlnLeuValProProAspGlnPheLysGlyThrPheArgTyrPaspGlnTyr 392
Db 333 TATTAATTTGAACTTGTTCGCGCAGATGAGATTCAAAGAACATTCAGATGGATGATAC 274
QY 393 LeuGlnLysSerAlaGlnLysThrLeuProLeuAspLeuPheLysPromeIleProSerGln 412
Db 273 TTGGAGAAAGAAATCTGACAAACCTTACCGCTTGACTTGTTCACCAATGCTTCCCA 214
QY 413 GlnArgLeuAspLysPheLysVal 420
Db 213 GACAGATTAGACAAATTTAAGGTA 190

RESULT 8
C42305 360 bp mRNA linear EST 18-OCT-1999
LOCUS C42305
DEFINITION C42305 Yuj1 Kohara unpublished cDNA:Strain N2 hermaphrodite embryo
```

Caenorhabditis elegans cDNA clone yk292d9 5', mRNA sequence.
ACCESSION C42305
VERSION C42305.1 GI:2378542
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Motolashi,T., Tabara,H., Watanabe,H., Sugimoto,A., Sano
1 (bases 1 to 360)
TITLE Expression map of the C.elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..360
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk292d9"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone.lib="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 119 a 62 c 78 g 101 t
ORIGIN
Alignment Scores:
Pred. No.: 3,76e-110 Length: 360
Score: 119.00 Matches: 119
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 23.90% Indels: 0
DB: 14 Gaps: 0
US-09-872-523-5 (1-498) x C42305 (1-360)
QY 9 ArgAlaasnLysLysSerAspArgLysLeuAspLysThrTyrLeuTrpGluSerTyrLeu 28
DB 2 AGAGCTAACCAAAAATCGACAGAAACTCGATAAGCTACTGTGGAAATCGTATTTA 61
QY 29 HisGlnPheGluLysGlyLysThrSerPheIleProValGluAlaPheAsnArgAsnLeu 48
DB 62 CATCAGTTCGAAAGGAAAACTTCTTCATTCAGTTGAAGCATTCATCGTAACCTT 121
QY 49 ThrValAsnPheAsnGluCysValLysGluGlyValIlePheGluThrValAlaHisAsp 68
DB 122 ACAGTTAATTTTAACGAATGCGTGAAGAGAGATTATCTCGAAACAGTGTCCATGAT 181
QY 69 TyrAspLysAsnGlySerPheIleGlnValArgTppPheAlaArgIleGluLysValCys 88
DB 182 TATGACAAGAACTCGCATTCGATCAAGTCACATGCTTGGACGAATGAAAAAGTTTGC 241
QY 89 GilyTyrArgValLeuAlaGlnPheIleGlyAlaAspThrLysPheTrpLeuAsnIleLeu 108
DB 242 GGATACAGAGTTCTGGCTCAGTTTATCGAGAGCTGACAGAAATTTGGCTCAATATTTTA 301
QY 109 SerAspAspMetPheGlyLeuAlaAsnAlaAlaMetSerAspProAsnMetAspLys 127
DB 302 TCGACGATATGTTTGGTTGGCAACGCCGAATGAGATCCCAATATGATATAA 358
RESULT 9
LOCUS AV197399 380 bp mRNA linear EST 26-JUL-1999
DEFINITION AV197399 Yuji Kohara unpublished cDNA:Strain N2 hermaphrodite
embryo Caenorhabditis elegans cDNA clone yk65h11 5', mRNA

sequence.
ACCESSION AV197399
VERSION AV197399.1 GI:5581170
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
REFERENCE Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
AUTHORS Kohara,Y., Shin,I.,T., Thierry-Mieg,J., Thierry-Mieg,D., Mitsuki,H.,
1 (bases 1 to 380)
TITLE Expressed genes in C.elegans
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
FEATURES
source 1..380
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk65h11"
/sex="hermaphrodite"
/dev_stage="embryo"
/clone.lib="Yuji Kohara unpublished cDNA:Strain N2
hermaphrodite embryo"
BASE COUNT 125 a 63 c 83 g 108 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 4,96e-103 Length: 380
Score: 112.00 Matches: 125
Percent Similarity: 99.21% Conservative: 0
Best Local Similarity: 99.21% Mismatches: 0
Query Match: 22.49% Indels: 1
DB: 9 Gaps: 0
US-09-872-523-5 (1-498) x AV197399 (1-380)
QY 6 LysIleValArgAlaAsnLysSerAspArgLysLeuAspLysThrTyrLeuTrp-GI 25
DB 3 AAAATGTCAGAGCTAACCAAAAATCGACAGAAACTCGATAAGACTTGTGGGA 62
QY 25 USerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIleProValGluAlaPheAs 45
DB 63 ATCTATTTTACATCGATTCGAGAAAGAAAACTTCTTCATTCAGTTGAAGCATTTCAA 122
QY 45 nArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyValIlePheGluThrVa 65
DB 123 TCGTAACCTTACAGTTAATTTTAACGAATGCGTGAAGAGAGATTATCTCGAAACAGT 182
QY 65 lValHisAspLysAsnGlySerPheIleGlnValArgTppPheAlaArgIleGlu 85
DB 183 GGTCCATGATATATGACAAGAACTCGATTCGATTAAGTCACATGCTTGGACGAATTGA 242
QY 85 uLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAspThrLysPheTrpLe 105
DB 243 AAAAGTTTGGGATACAGAGTTCTGGCTCAGTTTATCGAGCTGACAGAAATTTGGCT 302
QY 105 uAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaAlaMetSerAspProAsnMe 125
DB 303 CAATATTTTATCGACGATATGTTGTTGGCAACGCCGAATGAGATCCCAATAT 362
QY 125 tAspLysIleValTyr 130
DB 363 GGATAAAATTTGTATAT 378

RESULT 10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
AU216908/c	AU216908	unpublished oligo-capped cDNA library, stage L4	677 bp	linear	EST 17-JUL-2001							
	AU216908	Caenorhabditis elegans cDNA clone yk843e05 3', mRNA sequence.										
	AU216908	1 (bases 1 to 677)										
	AU216908.1	GI:14855065										
	EST.											
	Caenorhabditis elegans											
	Caenorhabditis elegans											
	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae											
	1 (bases 1 to 677)											
	kohara.Y., Shin-I., Thiery-Mieg, D., Thiery-Mieg, D., Suzuki, Y.											
	and Sugano, S.											
	A complementary view of the C.elegans genome											
	Unpublished											
	Contact: Yuji Kohara											
	Genome Biology Lab.											
	National Institute of Genetics											
	Yata 1111, Mishima, Shizuoka 411, Japan											
	Tel: 81-559-81-6854											
	Fax: 81-559-81-6855											
	Email: ykohara@lab.nig.ac.jp.											
	Location/Qualifiers											
	1. 677											
	/organism="Caenorhabditis elegans"											
	/mol_type="mRNA"											
	/strain="N2"											
	/db_xref="taxon:6239"											
	/clone="yk843e05"											
	/sex="Hermaphrodite"											
	/tissue_type="whole animal"											
	/dev_stage="L4"											
	/clone_lib="unpublished oligo-capped cDNA library, stage L4"											
	BASE COUNT	210 a	136 c	130 g	201 t							
	ORIGIN											
	Alignment Scores:											
	Pred. No.:	1.02e-101				Length:	677					
	Score:	111.00				Matches:	111					
	Percent Similarity:	100.00%				Conservative:	0					
	Best Local Similarity:	100.00%				Mismatches:	0					
	Query Match:	22.29%				Indels:	0					
	DB:	9				Gaps:	0					
	US-09-872-523-5 (1-498) x AU216908 (1-677)											
Oy	310	GLYGLINLSPHEGLULEULEILASPProleuNlaGInGInPheASnSLeuHISValAla	329									
Db	676	GGACAAAGTTTGAGCTCATCGACCCCTGGCTCAGCAATTCATTAACCTCCAGTGCCT	617									
Oy	330	SeTlleuLyuSPheCysLysTThGluGlyTyrLeuIleValGlyMetAspGlyProAsp	349									
Db	616	TCGATTTCCAATTTTGGCAAAACGAGGATATCTTATTGGGAGATGGATGCTCCAGAT	557									

RESULT 11	686 bp	mRNA	linear	EST 24-JAN-2002
LOCUS	BJ153087/c			
DEFINITION	BJ153087 unpublished oligo-capped cDNA library, C. elegans L1 stage			
ACCESSION	Caenorhabditis elegans cDNA clone yk1315b11 3', mRNA sequence.			
VERSION	BJ153087			
KEYWORDS	BJ153087.1 GI:18321072			
SOURCE	EST.			
ORGANISM	Caenorhabditis elegans			
REFERENCE	Caenorhabditis elegans			
AUTHORS	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae; 1 (bases 1 to 686)			
TITLE	1 (bases 1 to 686)			
JOURNAL	Kohara, Y., Shu, L.T., Thierry-Mieg, J., Thierry-Mieg, D., Suzuki, Y.			
COMMENT	and Sugano, S.			
	Unpublished			
	Contact: Tadashi Shin-i			
	Center For Genetic Resource Information			
	National Institute of Genetics			
	1111 Yata, Mishima, Shizuoka 411-8540, Japan			
	Tel: 81-559-81-6856			
	Fax: 81-559-81-6855			
	Email: tshini@genes.nig.ac.jp.			
FEATURES	Location/Qualifiers			
source	1..686			
	/organism="Caenorhabditis elegans"			
	/mol_type="mRNA"			
	/strain="N2"			
	/db_xref="taxon:6239"			
	/clone="yk1315b11"			
	/sex="hermaphrodite"			
	/tissue_type="whole animal"			
	/dev_stage="L1"			
	/clone_idb="unpublished oligo-capped cDNA library, C. elegans L1 stage"			
BASE COUNT	214 a 133 c 131 g 204 t 4 others			
ORIGIN				
Alignment Scores:				
Pred. No.:	1.13e-98	Length:	686	
Score:	108.00	Matches:	108	
Percent Similarity:	100.00%	Conservative:	0	
Best Local Similarity:	100.00%	Mismatches:	0	
Query Match:	21.69%	Indels:	0	
DB:	12	Gaps:	0	
US-09-872-523-5 (1-498) x BJ153087 (1-686)				
QY	313 PheGluLeuIleAspProLeuAlaGlnInPheAsnLeuHisValAlaSerIleLeu	332		
DB	684 TTGGAGCCATCGACCCMTTGCGCTCAGCAATTCATTACCTCCACGTCGCTTCGATCTC	625		
QY	333 LysPheCysLysThrGluGlyTyrLeuIleValGlyMetAspGlyProAspAlaLeuGlu	352		
DB	624 AAATTTTCCAAAACGAGAGATATCTTTGTGGGAAGATGTCGACGATGCACCTTCAA	565		
QY	333 AspSerPheProIleHisIleAsnAsnThrPheMetPheProValGlyTyrAlaGluLys	372		
DB	564 GACAGTTTTCCTATTCATATCATATATCATTTATGTCCACGATGGTTATGCGGAAAG	505		
QY	373 TyrAsnLeuGluLeuValProProAspLupheLysGlyThrPheArgTrpAspGluTyr	392		
DB	504 TATAATTTGGAACTGTGTCGCCGACATGATGTTCAAGAGAACATTCAGATGGATGAATAC	445		
QY	393 LeuGluLysGluSerAlaGluThrLeuProLeuAspLeuPheLysProMetProSerGln	412		
DB	444 TTGGAGAAAGATTCGCGAAACCTTACCGCTTACTGTTCAAGCCATGCTCTTCCAA	385		
QY	413 GluArgLeuAspLysPheLysVal 420			
DB	384 GAGAGATTGACCAATTTAAAGTA 361			

RESULT 12
A011588 340 bp mRNA linear EST 19-OCT-2000
LOCUS A011588 unpublished oligo-capped cDNA library Caenorhabditis
DEFINITION elegans cDNA clone yk7325 5', mRNA sequence.
ACCESSION A011588
VERSION A011588.1 GI:10925155
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 340)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1. 340
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk7325"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_11b="unpublished oligo-capped cDNA library"

BASE COUNT 111 a 58 c 72 g 98 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 5.88e-93 Length: 340
Score: 102.00 Matches: 102
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 20.48% Indels: 0
DB: Gaps: 9

US-09-872-523-5 (1-498) x A011588 (1-340)

OY 2 SerGIuPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLysThr 21
DB 34 TCTGAATTTCTGAAAATGTCAGACCTAAACAAAATCGACAGAAACTCGATTAAGACC 93
OY 22 TyrLeuTPGIsuSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIleProVal 41
DB 94 TACTGTGGAACTCTATTATCATCATGTCAGAAAGGAAAACCTTCTTCATTCACAGTT 153
OY 42 GluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluLysValLysGluLysIle 61
DB 154 GAAGCATTCACATCGTAACCTTACATTAATTTTAACGAATGCGCTGAGAAAGACATTATC 213
OY 62 PheGluThrValValHisAspTyrAspLysAsnCySaSpSerIleGlnValArgTrpPhe 81
DB 214 TTCGAACAGTGGTCATGATTAATGACAAAGAACTCGATTTCAGTTCAGATGCTTT 273
OY 82 AlaArgIleGluLysValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAspThr 101
DB 274 GCACCAATTGAAAGATTGGCGGATACAGATTCTGCTCAGATTATCGACCTGACACG 333
OY 102 LysPhe 103
DB 334 AAATTT 339

RESULT 13
A0116296/c 300 bp mRNA linear EST 19-OCT-2000
LOCUS A0116296 unpublished oligo-capped cDNA library Caenorhabditis
DEFINITION elegans cDNA clone yk743a4 3', mRNA sequence.
ACCESSION A0116296
VERSION A0116296.1 GI:10929863
KEYWORDS EST.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.

FEATURES
source
1. 300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_11b="unpublished oligo-capped cDNA library"

BASE COUNT 95 a 63 c 54 g 88 t
ORIGIN

Alignment Scores:
Pred. No.: 1.26e-62 Length: 300
Score: 72.00 Matches: 72
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.46% Indels: 0
DB: Gaps: 9

US-09-872-523-5 (1-498) x A0116296 (1-300)

OY 427 ValGIuLeuArgLeuGluAlaAlaAspMetCysGluAsnGlnPheIleCysProAlaThr 446
DB 299 GTAGGACGTACGCTTGAAAGCTGCTGACATGCTGAAATCACTTATTGTCACCTACA 240
OY 447 ValLysSerValHisGlyArgLeuIleAsnValAsnPheAspGlyTrpAspGluLuphe 466
DB 239 GTGAATTCAGTTCAGTGAAGACGTGATTAATGTCATTTGACGCGCTGGAGTCAACAATTT 180
OY 467 AspGluLeuTyrAspValAspSerHisAspIleLeuProIleGlyTrpCysGluAlaHis 486
DB 179 GATGAACGTATGATGATGAGTCCATCATATATTCACCGATRGATGATGATGACGCAC 120
OY 487 SerTyrValLeuGlnProProLysTyrTyrAsnTyr 498
DB 119 AGTTATGTTCTACCAACTCCGAAAAGTACACTAT 84

RESULT 14
A0115428/c 300 bp mRNA linear EST 19-OCT-2000
LOCUS A0115428 unpublished oligo-capped cDNA library Caenorhabditis
DEFINITION elegans cDNA clone yk7325 3', mRNA sequence.
ACCESSION A0115428
VERSION A0115428.1 GI:10928995
KEYWORDS EST.
SOURCE Caenorhabditis elegans

ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodermidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 300)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 300
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk732e5"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 94 a 60 c 54 g 89 t 3 others
ORIGIN

Alignment Scores:
Pred. No.: 1.34e-60 Length: 300
Score: 70.00 Matches: 70
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 14.06% Indels: 0
DB: 9 Gaps: 0

US-09-872-523-5 (1-498) x AU115428 (1-300)

QY 429 LeuAtgLeuGluAlaAlaAspMetCysGluAsnGlnPheIleCysProAlaThrValys 448
|||||
Db 298 CTACGCCCTGAGAGCTGCTACATGTGTGAATAATCATTTATTGTCTACACTACAGGANA 239
QY 449 SerValHisGlyArgLeuIleAsnValAsnPheAspGlyTrpAspGluLupheAspGlu 468
|||||
Db 238 TCAGTTCATGAGAGAGATGATTAATGTCAATTTGAGCGCTGGGATGAGAAATTTGATGAA 179
QY 469 LeuTyrAspValAspSerHisAspIleLeuProIleGlyTrpCysGluAlaHisSerTyr 488
|||||
Db 178 CTGTATGATGATGACTCCCATGATATTTACCGATAGATGATGTGAGCGCACAGTAT 119
QY 489 ValLeuGlnProProLysLysTyrAsnTyr 498
|||||
Db 118 GTTCTACACCTCCGAAAAAGTACAACTAT 89

RESULT 15
AU112450 337 bp mRNA linear EST 30-MAY-2003
LOCUS AU112450
DEFINITION AU112450 unpublished oligo-capped cDNA library Caenorhabditis
elegans cDNA clone yk743a4 5', mRNA sequence.
ACCESSION AU112450
VERSION AU112450.2 GI:31237498
KEYWORDS EST.
SOURCE
ORGANISM Caenorhabditis elegans
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea
; Rhabditidae; Pelodermidae; Caenorhabditis.
REFERENCE 1 (bases 1 to 337)
AUTHORS Kohara,Y., Shin-I,T., Thierry-Mieg,J., Thierry-Mieg,D., Suzuki,Y.
and Sugano,S.
TITLE A complementary view of the C. elegans genome
JOURNAL Unpublished
COMMENT On Oct 19, 2000 this sequence version replaced gi:10926017.

Contact: Yuji Kohara
Genome Biology Lab.
National Institute of Genetics
Yata 1111, Mishima, Shizuoka 411, Japan
Tel: 81-559-81-6854
Fax: 81-559-81-6855
Email: ykohara@lab.nig.ac.jp.
Location/Qualifiers
1. 337
/organism="Caenorhabditis elegans"
/mol_type="mRNA"
/strain="N2"
/db_xref="taxon:6239"
/clone="yk743a4"
/sex="Hermaphrodite"
/tissue_type="whole animal"
/dev_stage="varied"
/clone_lib="unpublished oligo-capped cDNA library"

BASE COUNT 107 a 59 c 74 g 92 t 5 others
ORIGIN

Alignment Scores:
Pred. No.: 1.97e-52 Length: 337
Score: 62.00 Matches: 62
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 12.45% Indels: 0
DB: 9 Gaps: 0

US-09-872-523-5 (1-498) x AU112450 (1-337)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysSerAspArgLysLeuAspLys 20
|||||
Db 23 ATGCTGGAATTTCTGAAATTTGTCAGAGCTAACCAAAAAATCGACAGAAAATCGATTAAG 82
QY 21 ThrTyrLeuTrpGluSerTyrIleuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
|||||
Db 83 ACCTACTGTGGGAATCTATTTACATCAGTTTCGAGAAAGAAAACCTCTTCATTTCCA 142
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluGlyVal 60
|||||
Db 143 GTTGAAGCATTCATGCTAATCCTTACAGTTAATTTTAACGAATGCCTGAAGAGAGATT 202
QY 61 IlePhe 62
|||||
Db 203 ATCTTC 208

Search completed: September 2, 2003, 23:10:26
Job time : 3076 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
---------------	-------	----------------	--------	----	----	-------------

Description
Caenorhabditis elegans
C. elegans mutant
C. elegans mutant

C. elegans lin-61

C. elegans lin-61Q

Plant microsatelli

Plant microsatellite
Plant microsatellite
Plant microsatellite
Plant microsatellite

Plant microsatelli

Plant microsatellites

Plant microsatelli

Plant microsatellite
Plant microsatellite

Plant microsatellites

Plant microsatellite

Plant microsatelli
Plant microsatelli
Plant microsatelli
Plant microsatelli
Plant microsatelli
Human GDP-mannose
DNA encoding novel
Human cDNA clone (c
Oligonucleotide (c
Oligonucleotide (c
Human secreted pro

Human TLE-2 gene.

```
Human TLE-2 gene.  
DNA encoding novel  
Kidney injury assoc  
Human cDNA sequenc  
Sputumretrovirus DN  
Stromal cell-deriv  
Human stromal cell
```

```
FT      /*tag= a
FT      /product= "LIN-61 protein"
XX
XX      W0200194545-A2.
XX
XX      13-DEC-2001.
XX
XX      01-JUN-2001; 2001WO-US17909.
XX
XX      02-JUN-2000; 2000US-208802P.
XX
XX      (MAST ) MASSACHUSETTS INST TECHNOLOGY.
XX
XX      Horvitz HR, Davison EM, Lu X;
XX
XX      MPI: 2002-401590/43.
XX
XX      P-PSDB; ABB78650.
XX
XX      New Caenorhabditis elegans nucleic acid involved in tumor suppressor
XX      pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
XX      diagnosing and treating cell proliferative diseases such as cancer
XX
XX      Claim 24; Page 74; 116pp; English.
XX
XX      The present invention describes a substantially pure nematode C. elegans
XX      (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
XX      (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
XX      (see ABB78650), where the polypeptides comprise at least 130, 110, 130
XX      contiguous amino acids of the 386, 322 and 498 amino acid sequences given
XX      in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
XX      cytosolic activity, and can be used in gene therapy. (I) is useful for
XX      modulating proliferation of a cell, and for identifying the compound that
XX      modulates cell proliferation. (I) can be used for diagnosing an animal
XX      (preferably, human) for the presence of the cell proliferation disease,
XX      or an increased chance of developing the disease, by measuring LIN-8,
XX      LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
XX      animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
XX      using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
XX      (I) is also useful for identifying the compound that modulates cell
XX      proliferation. (I) and its mutants are useful for understanding cell
XX      proliferative diseases including cancer, as well as in diagnosing and
XX      treating cell proliferative diseases. (I) is also useful for identifying
XX      tumor suppressors in other species such as mammals and may be used to
XX      identify therapeutic compounds. The present sequence encodes C. elegans
XX      LIN-61 from the present invention.
XX
XX      Sequence 1497 BP; 497 A; 272 C; 312 G; 416 T; 0 other:
XX
XX      Alignment Scores:
XX      Pred. No.: 0 Length: 1497
XX      Score: 498.00 Matches: 498
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 24 Gaps: 0
XX
XX      US-09-872-523-5 (1-498) x ABL52497 (1-1497)
XX
XX      1 Metsergiupheleuyllevalargalaansylusysersarpargylsleuasp 20
XX      1 ATGCTGAAATTTCTGAAAATTTGTGAGAGCTAACAAAAATCGCAGCAAGAACTCGATAG 60
XX
XX      21 ThyrtrleutrgiuserTyrleuHISGlnphegIuysgIyLysThrSerPheIlePro 40
XX      1 ACCGACTGTGGGAAATCTATTACATCAGTTTCGAGAAAGAAAAAATCTTTTCATTTCA 120
XX
XX      41 ValalualpheasnaIrgasnlleuthrValasnPheasnGluCysValIysGluGlyVal 60
XX      121 GTTGAAGGATTCATCGTAACCTTACAGTTAATTTTAAAGAAATGCGTGAAGAGGACTT 180
XX
XX      61 IlelphelgIuThrValIValHISAspTyrAspLysAsnCysAspSerIleGlnValArgrTrp 80
XX      181 ATCTTCGAACAGTGGTCCATGATTAATGACAAAGACTCGATTCGATTCAGTCAAGTGG 240
```

```
QY      81 PheAlaArgrIleGluIysValCysGlyTyrArgrValIleuAlaGlnPheIleGlyIleAsp 100
DB      241 TTTCACAGAAATGAAAAAGTTTCGGATACAGAGTTCTGCTCAGTTTATTCGAGAGCTGAC 300
QY      101 ThrlYsPheThrrPheIuasnIleleuserAspAspMetPheGlyIleuAlaAsnAlaIleMet 120
DB      301 ACGAAATTTTGGCTCAATATTTTATTCGAGAGATATGTTTGTTGGCAAAAGCCCAAGT 360
QY      121 SerAspProAsnMetAspIysIleValIleAlaProProIleuAlaIleasnGluIyTr 140
DB      361 AGTATCCCAATATGATATAAATTTGATATAGTCGCCGCTTGCAATCAACGAAGAATAC 420
QY      141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValIleGlnThr 160
DB      421 CAAATGATATGTGTAATTTATGTAATAATTTGATGATGATGCAAAATCGTCGGCCAAACT 480
QY      161 SerIleuserProIysPheAspGluGlyValAlaIleuIleuserLysHISArgPheIysVal 180
DB      481 TCGCTGCTCCAAAATTCGATGAAAGGAGAGCTCTCTTAAGCAAGATCGCTTCCAAAGTT 540
QY      181 GlyGlnArgrleuGluIleuIleuAsnTyrSerAsnSerThrGluIleArgrValAlaArgrIle 200
DB      541 GGACAAAGCTGTGAACTATTAAATTTATTCCAATTTCTACTGAAATACCGCTAGCGCAAT 600
QY      201 GlnGluIleCysGlyArgrArgrMetAsnValSerIleThrIysIysAspPheProGluSer 220
DB      601 CAAGAAATATGTGGACAGCAAGAAATGATATCTATCAAAAGAAAGACTTTCCGAAATCG 660
QY      221 LeuProAspAlaAspAspAspArgGlnValAlaPheSerSerGlySerGlnTyrTrpIleAsp 240
DB      661 CTTCACATGACAGATGACAGACAGACAGACAGCTTTATAGCTCGATGATCAATTTGGATAGAC 720
QY      241 GlnGlySerPhePheIlePheProValIleGlyPheAlaValAsnGlyTyrGlnIleuAsn 260
DB      721 GAGGAGAGCTTCTCATATTTCCGTGGATTGGACACAGCATGATGATCAACTAAAT 780
QY      261 AlaIysIysGluIyrrIleGlnHISThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
DB      781 GCGAAAAAGGAATATATGATGACACACAAATAATCTCAAGCAATAAAAAATGGAGAA 840
QY      281 AsnProArgrTyrAspSerAspAspValAlaPheAspGlnIleuAlaLysAspProIleAsp 300
DB      841 AATCCAGATATGACTCGACAGACGATCAATTTGATCAATATGCAAAAGATCCAAATGAT 900
QY      301 PrometIletrPargIysValIysValIleGlyIlysnPheGluIleuIleAspProIleuAla 320
DB      901 CCCATGATTTGGAGAAAAAGTTAAAGTTGGACAAAGTTTGAGCTCATCGAACCCCTGGCT 960
QY      321 GlnGlnPheAsnAsnIleuHISValAlaSerIleleuIysPheCysLysThrGluGlyTyr 340
DB      961 CAGCAATTCATTAACCTCCACGTCGCTTCGATTTCCAAATTTTCCAAAACCTGAAAGATAT 1020
QY      341 LeuIleValIleGlyMetAspGlyProAspAlaIleuGluAspSerPheProIleHISIleAsn 360
DB      1021 CTTATTTGGGAAATGAGATGTCAGATGCAATGCAATGCAAGCAAGTTTCTATTCATATCAAT 1080
QY      361 AsnThrPheMetPheProValIleGlyTyrAlaGluIysTyrAsnIleuGluIleuValIlePro 380
DB      1081 AATACATTTATATGTTCCAGATGTTATCGGAAAAAGATATATTTGGAAACTGTGTTCCGCA 1140
QY      381 AspGluPheLysGlyThrPheArgrTyrAspGluIyrrIleuGluIysGluIleuValIleThr 400
DB      1141 GATGAGTTCAAAAGAAATTCAGATGAGTGAATTAATCTTGGAGAAAGAAATGCGCAAAAC 1200
QY      401 LeuProIleuAspIleuPheIysProMetProSerGlnGluArgrIleuAspIysPheIysVal 420
DB      1201 CTACCGGCTGACTGTGTCACCAATGCTTTCCCAAGAGAGATATGACAAATTTTAAGGTA 1260
QY      421 IleuIleSerLysArgrValIleGlyLeuArgrleuGluAlaIleAspMetCysGluAsnGln 440
DB      1261 ATTCTGATTTCCAAACGGGTAGGACTACGCTTGAAGCTGCTGACATGTGTGAATAATCAG 1320
```

OY		441	PheleCySPROalathnVallySSerValHISglYargLeuileASnValAsnPheASP	460
Dd		1321	TTTTATTTCGCAGCTACAGTGAATCACTTCATGGAAGACAGATAAATGTCAATTTCGAC	1380
OY		461	GlyTrPaspGJugLupPheaspGLIeNtYrAspValasPSerHisAsPIleLeuProIle	480
Dd		1391	GGCTGGCATGAAGAATTTCATGCAACTGTATGATGTGCACTCCATGATATTTCTACCGCAT	1440
OY		481	GIYTTPCYSGlUAAlAHISSerTYrValLeuGlInProPolYSLYTYrAsnTYr	498
Dd		1441	GGATGGTGTAAGCGCACAGCTTATGTTCTACAACTCCGAAAAGTCAACTAT	1494
RESULT 2				
ABL52515	ID	ABL52515	standard; DNA; 1497 BP.	
XX	AC	ABL52515;		
XX	XX	17-JUL-2002	(first entry)	
DE	C.	elegans mutant lin-61	DNA sequence lin-61(n3446) SEQ ID NO:73.	
XX	Caenorhabditis elegans; lin-8; lin-56; lin-61;	tumour suppressor;		
KW	cell proliferation; nematode; cancer; mutant; gene; ds.			
XX	Caenorhabditis elegans.			
OS				
XX	Key	Location/Qualifiers		
FT	CDS	1..1497		
FT		/tag= a		
FT		/product= "LIN-61(n3446) protein"		
PN	MO200194545-A2.			
PD	13-DEC-2001.			
XX	01-JUN-2001;	2001WO-US17909.		
PR	02-JUN-2000;	2000US-208802P.		
PA	(MASI) MASSACHUSETTS INST TECHNOLOGY.			
PL	Horvitz HR,	Davison EM, Lu X;		
XX	WP1: 2002-401590/43.			
DR	P-PSDB; ABB78697.			
XX				
PT	New Caenorhabditis elegans nucleic acid involved in tumor suppressor			
PT	pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for			
PT	diagnosing and treating cell proliferative diseases such as cancer	-		
PS	Claim 27; Page 112-113; 116pp; English.			
XX				
XX	The present invention describes a substantially pure nematode C. elegans			
CC	(Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide			
CC	(see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide			
CC	(see ABB78650), where the polypeptides comprise at least 130, 110, 130			
CC	continuous amino acids of the 386, 322 and 498 amino acid sequences given			
CC	in ABB78648 to ABB78650, and modulate cell proliferation. (I) has			
CC	cytostatic activity, and can be used in gene therapy. (I) is useful for			
CC	modulating proliferation of a cell, and for identifying the compound that			
CC	(preferably, human) for the presence of the cell proliferation disease,			
CC	or an increased chance of developing the disease, by measuring lin-8,			
CC	lin-56 or lin-61 nucleic acid expression in a sample obtained from the			
CC	animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample			
CC	using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.			
CC	(I) is also useful for identifying the compound that modulates cell			
CC	proliferation. (I) and its mutants are useful for understanding cell			
CC	proliferative diseases including cancer, as well as in diagnosing and			
CC	treating cell proliferative diseases. (I) is also useful for identifying			
CC	tumour suppressors in other species such as mammals and may be used to			
CC	identify therapeutic compounds. The present sequence encodes a			

[illegible]

```

QY 301 prometilletrparglyvalylsvalglyclnlysphegileuleileasproleuala 320
    |||||||
Db 901 CCCATGATTTGGAGAAAGTTAAGTTGAGCAAAAGTTGAGCTCATCGACCCCTGGCT 960
QY 321 GlnGlnpnehsnlnleuhsvalAlaSerlleuleuysphecyslysthrngllytyr 340
    |||||||
Db 961 CAGCAATTCATTAACCTCCACGCTGCTTCGATTCCTCAAAATTTTCGAAACCTGGAAGGATAT 1020
QY 341 leuilevalglymetaspglyproaspalauleuGluaspserpheproillehsileasn 360
    |||||||
Db 1021 CTTATTTGGGAATGATGATGTCGAGATGCACCTTAACAACAGTTTCTCATTCATATCAAT 1080
QY 361 AsnThrphemetpneprovalglytyralagluulyslyrsnleuGluvalpropro 380
    |||||||
Db 1081 AATTCATTTATGTTCCACGATGGTGTATGCGGAAAAAGATATAATTTGGAACCTGTTCCGCCA 1140
QY 381 AspGluPhelysglythrpheargtrpaspGluThrleuGluysGluSerAlaGluThr 400
    |||||||
Db 1141 GATGAGTTCAAGAGCAATTCACATGGGATGATTAATCTTGGGAAAGAAATCTGCAGAAACC 1200
QY 401 LeuProleuaspLeuphelysPrometProSerGln-GluArgleuaspLyshelysVa 420
    |||||||
Db 1201 CTACCGCTGACTGTTTCAAGCCCAATGCCCTTCTTA-AGAGAGATTAGCAAAATTTAAAGT 1259
QY 420 lileleuileserlysarvalglyleuargleuGluAlaAlaspmetcysgluasngl 440
    |||||||
Db 1260 AATTCGATTTCCAAAGCGTAGGACTACGCTTGAAGCTGCTGACATGTTGTGAAATCA 1319
QY 440 nPheIlleCysProAlaThrValLysserValHisGlyArgleuileasnValaspheas 460
    |||||||
Db 1320 GTTTATTTGTCACACTACAGTAAGTAATCATGTTCAATGGAAGACTGATTAATGTCATTTGCA 1379
QY 460 pGlyTrpAspGluLupheaspGluLeuTyraSpValaspSerHisaspIleuProIl 480
    |||||||
Db 1380 CGGCTGGGATGAAGAATTTGATGAACGTATGATGTGATGCCATCATGATTTTACCCAT 1439
QY 480 eGlyTrpCysGluAlaHisSerlyrvalleuGlnProProCyslystyrAsnlyr 498
    |||||||
Db 1440 AGGATGGTGTAACGCGCACAGTTATGTTCTACAACTCCGAAAAAGTACAACTAT 1494

```

```

DR WPI: 2002-401590/43.
DR P-PSDE: AB878698.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Claim 27; Page 113; 116pp; English.
XX
CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see AB878648), LIN-56 polypeptide (see AB878649), or LIN-61 polypeptide
CC (see AB878650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in AB878648 to AB878650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring LIN-8,
CC LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence encodes a
CC C. elegans mutant LIN-61 protein from the present invention.
XX
SQ Sequence 1497 BP; 498 A; 272 C; 311 G; 416 T; 0 other;

Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 411.00 Matches: 497
Percent Similarity: 99.60% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 82.53% Indels: 2
DB: 24 Gaps: 0

US-09-872-523-5 (1-498) x ABL52516 (1-1497)
QY 1 MetSerGluPhelysIleValArgAlaAsnLysserAspArglylsleuAspLys 20
    |||||||
Db 1 ATGTCTGATTTTCGAAATTTGTCAGACCTTAACAAAATCGCAGAAAACCTGATGAG 60
QY 21 ThrTyrlleuTrpGluSerTyrlleuHisGlnPheGluLysglyThrSerPheIlePro 40
    |||||||
Db 61 ACCTACTGTGGGAATCCTATTATACATCAGTTCGAGAAAGAAAACCTTCTTCATTCGA 120
QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysVallysgluGlyVal 60
    |||||||
Db 121 GTTGAGACATTCACATCGTAACCTTACAGTAAATTTTAAAGAAATCGCTGAAGAGACTT 180
QY 61 llePheGluThrValIleValHisAspTyraSpLysAsnCysAspSerlleGlnValArgTrp 80
    |||||||
Db 181 ATCTTCGAAACAGTGTGTCATGATTAACAGAACTCGATTCGATTCAGATCAGATGG 240
QY 81 PheAlaArgIleGluLysValCysGlyTyraArgValleuAlaGlnPheIleGlyAlaAsp 100
    |||||||
Db 241 TTTCGACGAATTTGAAAAAGTTTGGCGGATACAGAGTTGCTGCAATTTATGAGAGCTGAC 300
QY 101 ThrlyspheTrpPleuAsnIleleuSeraspMetPheGlyleuAlaAsnAlaIleMet 120
    |||||||
Db 301 ACGAAATTTTGGCTCAAAATTTTATCGGACATATGTTGGTTGGCAAAAGCGCAATG 360
QY 121 SeraspProAsnMetAspLysIleValTyraIaProProleuAlaIleasnGluGluTyra 140
    |||||||
Db 361 AGTGAATCCCAATATGATGAATAATTTGATATGCTCCGCCGCTTGCATTAAGAGAAATAC 420
QY 141 GlnAsnAspMetValAsnTyraValAsnAsnCysIleaspGlyGluIleValGlyGlnThr 160
    |||||||

```

```

Db      421 CAAATGATATGTAATTATGTAAATAATTCATTGATGCGCAAACTCGGCCAAACT 480
Qy      161 SerLeuSerProLysPheAspGluLysAlaLeuSerLysHisArgPheLysVal 180
Db      481 TCGCTGCTCCAAATTCATGAGGAAAGGCTCTTAAGCAAGATGCTTCAAGTT 540
Qy      181 GluGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluLeuArgAlaArgIle 200
Db      541 GGCACAAAGCTTGAACATTAATTAATTCACATTCACGAATACGCCGTACGCCAATT 600
Qy      201 GlnGluIleCysGluArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
Db      601 CAAGAAATATGTGAGACGAGAAATGAATGTATCATCCAAAGAAACCTTCCCGAATCG 660
Qy      221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpPheAsp 240
Db      661 CTTCGAGATGCAGATGACACACACAGCAAGCTTTAGCTGAGCTCAATATTTGGATGAC 720
Qy      241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
Db      721 GAGGGAAGCTCTCTCATATTCCTGTTGATTTGCAGCATCAATGATATCACTAAAT 780
Qy      261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
Db      781 GCGAAAAAGGAATATATATGAGCACACAAATAAATTTGCTCAAGCAATAAATAATGAGAA 840
Qy      281 AsnProArgTyrAspSerAspAspValThrPheAspGluLeuAlaLysAspProIleAsp 300
Db      841 AATCCAAAGATATGACCTCACACGACGTCACATTTGATCAATTACCAAAATATCCAAATTGAT 900
Qy      301 PrometIleTyrArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
Db      901 CCCATGATTTGGAGAAAGTTAAAGTTGGCAAAAGTTTACGCTCATCCACCCCTTGCGCT 960
Qy      321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
Db      961 CAGCAATTCATATTAACCTCCACGCTCGCTGATTTCTCAAAATTTTGGCAAACTGAGGATAT 1020
Qy      341 LeuIleValGlyMetAspGlyProAspAlaLeuGluAspSer-PheProIleHisIleAs 360
Db      1021 CTTATGTTGGAAATGATGTCACAGATGCACATTGAAGACAA-TTTTCTTATTCATATCA 1079
Qy      360 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPr 380
Db      1080 TATACATTTATGTTCCAGTTGCTTATGCGGAAAGTAAATTTGCAACTTGTTCGCC 1139
Qy      380 AspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaLuh 400
Db      1140 AGATGAGTTCAAGCAACATTCAGATGGATGAATCTTGAGAAAGAAATCTGCAGAAAC 1199
Qy      400 rLeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVa 420
Db      1200 CCGACGCTTGAACCTTTCACAGCAATGCTTCCCAAGAGATTTGACAAATTTAAGT 1259
Qy      420 IleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnG 440
Db      1260 AATTCGATTTCCAAACGGGTAGGACTACGCCCTTGAAGCTGCTGACATGTGTGAATAAT 1319
Qy      440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPhas 460
Db      1320 GTTATTTTCCAGCTACAGTGAATCACTTCATGGAAACATGATTAAGTCAATTTGCA 1379
Qy      460 pGlyTrpAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI 480
Db      1380 CCGCTGGATGGAAGATTTGATGAACTGATGATGATGACTCCCAATATTTCTACCGAT 1439
Qy      480 eGlyTrpCysGluAlaHisSerTyrValLeuGlnProPolysLysTyrAsnTyr 498
Db      1440 AGATATGTTGGAAGCGCACAGTTATGTCTACAACCTCGGAAAAAGTACAACATAT 1494

```

RESULT 4
ABLS2517

ID ABL52517 standard; DNA; 1497 BP.

```

XX      XX      ABL52517;
AC      XX      17-JUL-2002 (first entry)
DT      XX      C. elegans mutant lin-61 DNA sequence lin-61(n3624) SEQ ID NO:75.
DE      XX      Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KW      XX      cell proliferation; nematode; cancer; mutant; gene; ds.
OS      XX      Caenorhabditis elegans.
FH      XX      Key Location/Qualifiers
FT      FT      CDS 1..1497
FT      FT      /tag= a
FT      FT      /product= "LIN-61(n3624) protein"
PN      PN      WO200194545-A2.
PD      PD      13-DEC-2001.
PE      PE      01-JUN-2001: 2001MO-US17909.
PR      PR      02-JUN-2000: 2000US-208802P.
PA      PA      (MAS1 ) MASSACHUSETTS INST TECHNOLOGY.
PI      PI      Horvitz HR, Davison EM, Lu X;
DR      DR      WPI; 2002-401590/43.
XX      XX      P-PSDB; ABB78659.
PT      PT      New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT      PT      pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT      PT      diagnosing and treating cell proliferative diseases such as cancer
XX      XX      Claim 27; Page 113-114; 116pp; English.
XX      XX      The present invention describes a substantially pure nematode C. elegans
XX      XX      (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
XX      XX      (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
XX      XX      (see ABB78650), where the polypeptides comprise at least 130, 110, 130
XX      XX      contiguous amino acids of the 386, 322 and 498 amino acid sequences given
XX      XX      in ABB78648 to ABB78650, and modulate cell proliferation. (I) has
XX      XX      cytostatic activity, and can be used in gene therapy. (I) is useful for
XX      XX      modulating proliferation of a cell, and for identifying the compound that
XX      XX      modulates cell proliferation. (I) can be used for diagnosing an animal
XX      XX      (preferably, human) for the presence of the cell proliferation disease,
XX      XX      or an increased chance of developing the disease, by measuring lin-8,
XX      XX      lin-56 or lin-61 nucleic acid expression in a sample obtained from the
XX      XX      animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample
XX      XX      using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
XX      XX      (I) is also useful for identifying the compound that modulates cell
XX      XX      proliferation. (I) and its mutants are useful for understanding cell
XX      XX      proliferative diseases including cancer, as well as in diagnosing and
XX      XX      treating cell proliferative diseases. (I) is also useful for identifying
XX      XX      CC tumour suppressors in other species such as mammals and may be used to
XX      XX      CC identify therapeutic compounds. The present sequence encodes a
XX      XX      C. elegans mutant LIN-61 protein from the present invention.
XX      XX      Sequence 1497 BP; 497 A; 271 C; 312 G; 417 T; 0 other;
SQ      SQ      Alignment Scores:
XX      XX      Pred. No.: 0 Length: 1497
XX      XX      Score: 397.00 Matches: 497
XX      XX      Percent Similarity: 99.608 Conservative: 0
XX      XX      Best Local Similarity: 99.608 Mismatches: 1
XX      XX      Query Match: 79.728 Indels: 2
XX      XX      DB: 24 Gaps: 0
US-09-872-523-5 (1-498) x ABL52517 (1-1497)
Qy      Qy      1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20

```

```

Db      1 ATGCTCAATTTTCGAAAAATTTGAGCTAACAAAAATCGACAAAAAATCTGGATGAG 60
Oy      21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluTyrSglYsthrSerPheIlePro 40
Db      61 ACCACTTGTGGGAAATCTTATTTACATCAGTTCGAGAAAGAAAAAATCTTCTTATTCGA 120
Oy      41 ValGluAlaPheAsnArgAsnLeuThrValAsnDheAsnGluCysValLysGluGlyVal 60
Db      121 GTTGAACCATTTCAATCGTAACTTACAGTTAAATTTTAAAGCAATCGTGAAGAGAGAGTT 180
Oy      61 IllePheGluTrpThrValHisAspTyrAspLysAsnGlySerIleGlnValArgTyr 80
Db      181 ATCTTCGAAACAGTGGCCATGATTATGACAAAGATCGATTCGATTCAAGTCAGATGG 240
Oy      81 PheAlaArgIleGluTyrValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAsp 100
Db      241 TTTTCAGCAATTTGAAAAAGTTTGGCGATACAGACTTCGGCTCAGTTTATCGAGCTGAC 300
Oy      101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaLamet 120
Db      301 ACGAAATTTTGGCTCAATATTTTATCGGACGATATGTTGGTTGGCAAAACGCCGCATG 360
Oy      121 SerAspProAsnMetAspLysIleValTyrAlaPro-ProIleuAlaIleAsnGluGly 140
Db      361 AGTATCCCAATATGAGATMAAATGTATATGCTTTC -GCCGCTTGCAATCAACGAGAAATA 419
Oy      140 rGlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnTh 160
Db      420 CCAAAATATATGTGTAATTTATGTAATAATTTGCATTGATGGCAAAATCGTCGGCAAAAC 479
Oy      160 rSerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVa 180
Db      480 TTCCGTCTCTCCAAATTCGATGAAAGGAGCGCTCTCTTAAGCAACGCTTTCCTCAAGT 539
Oy      180 IGLYlnArpLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgI 200
Db      540 TGGACAAAGCTCTTGAACTATTAATTTATCCAAATTTCTGAATATCCCGTAGCCGCAAT 599
Oy      200 eGlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSe 220
Db      600 TCAAGAAATATGTGGACGACGACGAATGATGTATCTATCAAAAGAAAGACTTCCCGCAATC 659
Oy      220 rLeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAs 240
Db      660 GCTTCCAGATGATGATGACGACGACGACCAAGCTTTAGCTCTGATCTCAATATTTGGATAGA 719
Oy      240 pGluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAs 260
Db      720 CGAGGGAAGCTTCTTCATATTTCCGTGGATTTGCACAGCATGATGATACACTATAA 779
Oy      260 nAlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyG 280
Db      780 TGGCAAAAAGGAATATATTTAGACACACAAATTAATTTGCTCAAGCAATAAAAAAGGAGA 839
Oy      280 uAsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAs 300
Db      840 AAATCCAAAGATATGATCAGACGACGATCACAATTTATGATCAATTAAGAAAAAGATCAATGA 899
Oy      300 pProMetIleTrpAlaGlyValLysValGlyGlnLysPheGluLeuIleAspProLeuAl 320
Db      900 TCCCATGATTTTGGACAAAAAGTTAAGCTTGGACAAAAAGTTTGAAGCTCATCGACCCCTTGGC 959
Oy      320 aGlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
Db      960 TCAGCAATTCATTAACCTCCACGCTCGCTTCGATCTCAAAATTTTCCAAAAAGAGAGATA 1019
Oy      340 rLeuIleValGlyMetAspGlyProAspAlaLeuLysAspSerPheProIleHisIleAs 360
Db      1020 TCTTATTTGGGAAATGGATGGTCCAGATGCACCTTGAAGCAACATTTTCCATATTCATATCA 1079
Oy      360 nAsnThrPheMetPheProValGlyTyrAlaGluTyrTrpAsnLeuGluLeuValProPr 380

```

```

Db      1080 TAATACATTTATGTTCCCACTTGGTTATGGGAAAGATATAATTTGGAACTGTTCGCC 1139
Oy      380 oAspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaGluTh 400
Db      1140 AGATGAGCTTCAAAAGAACATTTCAAGATGGATGATATCTTGAGAAAAGAAATCTCGAGAAAC 1199
Oy      400 rLeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVa 420
Db      1200 CCTACCGCTTGACTGTTCACAGCCAAATGCCCTTCCCAAGAGAGATTAGACAAATTTAAGCT 1259
Oy      420 lIleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnG 440
Db      1260 AATTCGATTTCCAAACGGGAGAGACTACGCTTGAAGCTGCTGACATGTGTGAAAAATCA 1319
Oy      440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
Db      1320 GTTATTTTGTTCACACTCAGAGAAATAGTTCAAGGAAGACTGATTAATGTCAATTTTCA 1379
Oy      460 pGlyTrpAspGluLubPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI 480
Db      1380 CGGCTGGAGTGAAGAAATTTGATGAACTGTATGATGTGACTCCCATGATATTTTACCGAT 1439
Oy      480 eGlyTrpCysGluAlaHisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
Db      1440 AGATGGTGTGAACGACGACAGTTATGTCTCAACCTCCGAAAAAGTACAACTAT 1494

RESULT 5
ABL52518 standard; DNA; 2307 BP.
XX ABL52518;
AC ABL52518;
XX 17-JUL-2002 (first entry)
DE C. elegans lin-61 DNA sequence SEQ ID NO:76.
XX Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;
KW cell proliferation; nematode; cancer; mutant; gene; ds.
XX Caenorhabditis elegans.
OS WO200194545-A2.
XX 13-DEC-2001.
XX 01-JUN-2001; 2001WO-US17909.
XX 02-JUN-2000; 2000US-208802P.
PA (MASI ) MASSACHUSETTS INST TECHNOLOGY.
PI Horvitz HR, Davison EM, Lu X;
DR WPI; 2002-401590/43.
XX New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Disclosure; Page 114-115; 116pp; English.
XX The present invention describes a substantially pure nematode C. elegans
XX (Caenorhabditis elegans) nucleic acid (i) encoding a LIN-8 polypeptide
XX (see ABB78648), LIN-56 polypeptide (see ABB78649), or LIN-61 polypeptide
XX (see ABB78650), where the polypeptides comprise at least 130, 110, 130
XX contiguous amino acids of the 386, 322 and 498 amino acid sequences given
XX in ABB78648 to ABB78650, and modulate cell proliferation. (i) has
XX cytosolic activity, and can be used in gene therapy. (i) is useful for
XX modulating proliferation of a cell, and for identifying the compound that
XX modulates cell proliferation. (i) can be used for diagnosing an animal
XX (preferably, human) for the presence of the cell proliferation disease,
XX or an increased chance of developing the disease, by measuring lin-8,
XX lin-56 or lin-61 nucleic acid expression in a sample obtained from the

```

CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61 DNA sequence from the present invention.

SO Sequence 2307 BP; 740 A; 393 C; 466 G; 708 T; 0 other:

Alignment Scores:

Pred. No.:	5.21e-285	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	24	Gaps:	0

US-09-872-523-5 (1-498) x ABL52518 (1-2307)

QY 150 AaNCysIIlaSpGlyluIleValIGlyInThrSerLeuSerProLysPheAspGluGly 169
DB 1138 AATTCATTGATGGCAAAATCGTCGGCAAACTCGCTGCCAAATTCGATGAAGG 1197
QY 170 LysAlaLeuLeuSerLysHsArGpHeLysValIGlyInArGluLeuLeuAaNTyr 189
DB 1198 AAGGCTCTCTTAAGCAAGCATCGTTCCAAAGTTGGCAACGCTTAACTATTAAATTAT 1257
QY 190 SerAsnSerThrGluIleArGyValAlaArGIIeGluIleCysGlyArGArGMeLaaN 209
DB 1258 TCCAATTCTACGAAATACGCGTACGCGCAATTCAGAAATATGTGACAGACAAATGAAAT 1317
QY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgIn 229
DB 1318 GATCTCATCAAAAGAAACCTTCCGAAATCGCTTCAGATGACATGACGACAGACAA 1377
QY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPheHeIlePheProVal 249
DB 1378 GTCCTTAGCTTCGATCTCAATATTTGGATAGACGAGGAAGCTTCTCATATTTCTCTGTT 1437
QY 250 GlyPheAlaAlaValaAsnGlyTyrGluLeuAsnAlaLysGluTyrIleGluHisThr 269
DB 1438 GGATTTGCAGACAGTCATGATGATCAACTAAATGCAAAAGAAATATATTTGAGCAGACA 1497
QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyGluAsnProArGTrAspSerAspVal 289
DB 1498 AATAAATTCCTCAACCAATATAAAATGAGAAATCCAAAGATATGACTCAGACGACGTC 1557
QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspProMetIleTrpArgLysValLysVal 309
DB 1558 ACATTTGATCAATTTACCAAAAGATCCCAATTTGATCCCATGTTTGGAGAAAGTTAAGGTT 1617
QY 310 GlyGlnLysPheGluLeuLeuLeaSPProLeuAlaGlnGlnPheAsnAsnLeuHisVala 329
DB 1618 GGCAGAAAGTTTGAGCTCATCGACCCCTTGGCTCAGACCAATTCATTAACCTCCACGTCGT 1677
QY 330 SerIleLeuLysPheCysLysThrGluGlyTyrLeuIleValIGlyMetAspGlyProAsp 349
DB 1678 TCGATTCCTCAAAATTTGCAAAACTGGAAGATATCTTATGTGGGAATGATGATCCAGAT 1737
QY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValIGlyTyr 369
DB 1738 GCACCTTGAAGACAGTTTCTCTATTCATATCAATTAATATTTATGTCCAGTTGGTAT 1797
QY 370 AlaGluLysTyrAsnLeuGluLeuValProProAspGluPheLysGlyThrPheArGTrp 389
DB 1798 GCGGAAAGATATTAATTTGGAACCTGTCCGCCAGATGAGTTCAAAGCAATTCACATG 1857
QY 390 AspGluTyrLeuGluLysGluSerAlaGluThrLeuProLeuAspLeuPheLysProMet 409
DB 1858 GATGATTAAGTGGAGAAAGATCTGCAGAAACCTTCGCTTGACTTGTTCAGCAAGT 1917

QY 410 ProSerGluIuArGLeuAspLysPheLysValIleLeuIleSerLysArGVal 427
DB 1918 CCTTCCCAAGAGATTAACAAATTTAAGTATTTCTATTTCCAAACGGTT 1971

RESULT 6
ID ABL52519 standard; DNA; 2307 BP.
AC ABL52519;
XX

DT 17-JUL-2002 (first entry)

DE C. elegans lin-61(sy223) DNA sequence SEQ ID NO:77.

KW Caenorhabditis elegans; lin-8; lin-56; lin-61; tumour suppressor;

KM cell proliferation; nematode; cancer; mutant; gene; ds.

OS Caenorhabditis elegans.

PN MO200194545-A2.

PD 13-DEC-2001.

PF 01-JUN-2001; 2001WO-US17909.

PR 02-JUN-2000; 2000US-208802P.

PA (MAST) MASSACHUSETTS INST TECHNOLOGY.

PI Horvitz HR, Davison EM, Lu X;

DR WPI; 2002-401590/43.

PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer

PS Disclosure; Page 115; 116pp; English.

XX The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see ABB78648), LIN-56 polypeptide (see ABB76649), or LIN-61 polypeptide
CC (see ABB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in ABB76648 to ABB78650, and modulate cell proliferation. (I) has
CC cytosolic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring lin-8,
CC lin-56 or lin-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC using Ab or measuring amount of lin-8, lin-56, lin-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(sy223) DNA sequence from the present invention.

SO Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other:

Alignment Scores:

Pred. No.:	5.21e-285	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	24	Gaps:	0

US-09-872-523-5 (1-498) x ABL52519 (1-2307)


```

OY 150 AsnCysIIleaspglygluilevalglycInThrSerLeuSerProLysPheaspglygly 169
    |||||
DB 1138 AATTGCATGTATGGCGAAATCGTCGGCCAAACTCGCTGCTCCAAAATTCGATGAAGGG 1197
OY 170 LysAlaLeuLeuSerLysHisArgPheLysValglyInArgLeuGluLeuLeuAsnTyr 189
    |||||
DB 1198 AAGGCTCTCTTAACCAAGCATCGTTCAAGTTGGACAGCTCTTGAATTAATTAAT 1237
OY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
    |||||
DB 1258 TCCAAATTCATCTGAATAATACGGTACCGCGAATTCAGAAATATGTGGACGAGATGAT 1317
OY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgGln 229
    |||||
DB 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTCCAGATGACAGATGACAGACAA 1377
OY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPhePheIlePheProVal 249
    |||||
DB 1378 GTCTTTAGCTCTGGATCTCAATATTTGGATAGACAGAGGAGACCTTCTCATATTTCTGTT 1437
OY 250 GlyPheAlaAlaValAsnGlyTyrGlnLeuAsnAlaLysLysGlyTyrIleGluHisThr 269
    |||||
DB 1438 GGATTTGGACGACGATCATGATATCACTAAATGCGAAAGAAATATATTGACGACACCA 1497
OY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyLysAsnProAlaGlyTyrAspSerAspAspVal 289
    |||||
DB 1498 AATAAAATGTCTCAAGCAATATAAATGAGAGAAATCCAAAGATATGACTCAGACGACGTC 1557
OY 290 ThrPheAspGlnLeuAlaLysAspProIleAspPheMetIleTyrPargLysValLysVal 309
    |||||
DB 1558 ACATTTGATCATTAATGCAAAAGATCCCAATTCATCCCATGATTGGAGAAAGTTAAGTT 1617
OY 310 GlyGlnLysPheGluLeuIleAspProLeuAlaGlnGlnPheAsnAsnLeuHisValAla 329
    |||||
DB 1618 GCAGCAAAAGTTGAGCTCATGACGCCCTGGCTCAGCAATTAACCTCCAGCTCCGCT 1677
OY 330 SerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProAsp 349
    |||||
DB 1678 TCGATTCCTCAAAATTTGCAAAAGTGAAGATATCTTATTTGGGATGATGATGCTCCAGAT 1737
OY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValAlaGlyTyr 369
    |||||
DB 1738 GCACCTGGAAGACAGTTTCCATTCATATCAATATATACATTTATGTTCCAGTTGTTAT 1797
OY 370 AlaGluLysTyrAsnLeuGluLeuValProProAspGluPheLysGlyThrPheArgTrp 389
    |||||
DB 1798 GCGGAAAGATTAATTTGGAACCTGTGTCGCCAGATGAGTTCAAAGGAACATTCAGATGG 1857
OY 390 AspGluTyrLeuGluLysGluSerAlaGluThrLeuProLeuAspLeuPheLysProMet 409
    |||||
DB 1858 GATGAAATCTTGAGAAAGAAATCTGCAGAAACCCCTACCGCTTGACTTGTTCAGCCCAATG 1917
OY 410 ProSerGlnGluArgLeuAspLysPheLysValIleLeuIleSerLysArgVal 427
    |||||
DB 1918 CCTTCCCAAGAGAGATTTAGACAAATTTAAGTAATTCGATTTCCAAACGGGTT 1971

```

```

XX 13-DEC-2001.
PD
XX
XX 01-JUN-2001; 2001MO-US17909.
PF
XX
PR 02-JUN-2000; 2000US-208802P.
PA
XX (MAST ) MASSACHUSETTS INST TECHNOLOGY.
XX
PI Horvitz HR, Davison EM, Lu X;
PI WPI; 2002-401590/43.
XX
PT New Caenorhabditis elegans nucleic acid involved in tumor suppressor
PT pathway, encoding LIN-8, LIN-56 or LIN-61 polypeptide, useful for
PT diagnosing and treating cell proliferative diseases such as cancer
XX
PS Disclosure; Page 116; 116pp; English.
XX
CC The present invention describes a substantially pure nematode C. elegans
CC (Caenorhabditis elegans) nucleic acid (I) encoding a LIN-8 polypeptide
CC (see AB78648), LIN-56 polypeptide (see AB78649), or LIN-61 polypeptide
CC (see AB78650), where the polypeptides comprise at least 130, 110, 130
CC contiguous amino acids of the 386, 322 and 498 amino acid sequences given
CC in AB78648 to AB78650, and modulate cell proliferation. (I) has
CC cyrostatic activity, and can be used in gene therapy. (I) is useful for
CC modulating proliferation of a cell, and for identifying the compound that
CC modulates cell proliferation. (I) can be used for diagnosing an animal
CC (preferably, human) for the presence of the cell proliferation disease,
CC or an increased chance of developing the disease, by measuring LIN-8,
CC LIN-56 or LIN-61 nucleic acid expression in a sample obtained from the
CC animal, by measuring LIN-8, LIN-56 or LIN-61 polypeptide in the sample.
CC using Ab or measuring amount of LIN-8, LIN-56, LIN-61 mRNA in the sample.
CC (I) is also useful for identifying the compound that modulates cell
CC proliferation. (I) and its mutants are useful for understanding cell
CC proliferative diseases including cancer, as well as in diagnosing and
CC treating cell proliferative diseases. (I) is also useful for identifying
CC tumour suppressors in other species such as mammals and may be used to
CC identify therapeutic compounds. The present sequence represents a
CC C. elegans lin-61(n3635) DNA sequence from the present invention.
XX
SO Sequence 2307 BP; 741 A; 393 C; 465 G; 708 T; 0 other;
XX
XX
Alignment Scores:
Pred. No.: 5,21e-285 Length: 2307
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.82% Indels: 0
DB: 24 Gaps: 0
US-09-872-523-5 (1-498) x ABL52520 (1-2307)
OY 150 AsnCysIIleaspglygluilevalglycInThrSerLeuSerProLysPheaspglygly 169
    |||||
DB 1138 AATTGCATGTATGGCGAAATCGTCGGCCAAACTCGCTGCTCCAAAATTCGATGAAGGG 1197
OY 170 LysAlaLeuLeuSerLysHisArgPheLysValglyInArgLeuGluLeuLeuAsnTyr 189
    |||||
DB 1198 AAGGCTCTCTTAACCAAGCATCGTTCAAGTTGGACAGCTCTTGAATTAATTAAT 1237
OY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
    |||||
DB 1258 TCCAAATTCATCTGAATAATACGGTACCGCGAATTCAGAAATATGTGGACGAGATGAT 1317
OY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgGln 229
    |||||
DB 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTCCAGATGACAGATGACAGACAA 1377
OY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPhePheIlePheProVal 249
    |||||
DB 1378 GTCTTTAGCTCTGGATCTCAATATTTGGATAGACAGAGGAGACCTTCTCATATTTCTGTT 1437

```


QY 250 GlyPheAlaAlaValaAsnGlyTyrGlnLeuAsnAlaLysGlyTyrIleGlnHisThr 269
 |||||||
 Db 1438 GGATTGGACGACGTCAATGATATCAAAATGCGAAAGAAATATTTGACGACACA 1497
 QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyLysAsnProAlaTyrAspSerAspVal 289
 |||||||
 Db 1498 AATAAATGGCTCAAGCAATAAATAATGAGAAATCCAAAGATATGACACGACGC 1557
 QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspPromeIleTyrArgLysVal 309
 |||||||
 Db 1558 ACATTGATCAATTCACAAACATCCAAATTCATCCATATTTGGAGAAAGTTAAGGTT 1617
 QY 310 GlnLysPheGlnLeuIleAspProLeuAlaGlnGlnPheAsnAsnLeuHisVala 329
 |||||||
 Db 1618 GGACAAAGTTTGAGCTCATGCGACCCCTTGGCTCAGCAATTCATTAACCTCCACGTCGCT 1677
 QY 330 SerIleLeuLysPheCysLysThrGlnGlyTyrLeuIleValGlyMetAspGlyProAsp 349
 |||||||
 Db 1678 TCGAATCTCAAAATTTGGCAAACTGAAAGATATCTTATGTGGAAATGATGTCGACAT 1737
 QY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValGlyTyr 369
 |||||||
 Db 1738 GCACCTTGAGACAGATTTCCTCATCATCAATTAATACATTAATGTTCCACGTTGGTTAT 1797
 QY 370 AlaGlnLysTyrAsnLeuGlnLeuValProAspGluPheLysGlyThrPheArgTyr 389
 |||||||
 Db 1798 GCGGAAAGTATATTAATTTGSACTTGTTCCGCGCAGATGATTCAAAGAACATTCAGATGG 1857
 QY 390 AspGlnTyrLeuGlnLysGlnSerAlaGlnThrLeuProLeuAspLeuPheLysProMet 409
 |||||||
 Db 1858 GATGATACTCTGAGAGAAAGAACTGCGAAGAACCTACCCGCTTGACTGTTCAAGCCAAATG 1917
 QY 410 ProSerGlnGlnArgLeuAspLysPheLysValIleLeuIleSerLysArgVal 427
 |||||||
 Db 1918 CCTTCCCAAGAGAGATTGACAAATTTAAGTAATTCGATTCCAAACGGGTT 1971

RESULT 8
 AAA31397/c
 ID AAA31397 standard; DNA; 344 BP.
 XX
 XX AAA31397;
 XX
 XX 05-JUL-2000 (first entry)
 DT
 XX
 XX Plant microsatellite marker #358.
 DE
 XX
 XX Plant microsatellite sequence; core repeat sequence: detection; probe;
 KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
 KW variety identification; genetic variability evaluation; primer; ss.
 XX
 XX Eucalyptus grandis.
 OS
 XX
 XX W09967421-A1.
 PN
 XX
 XX 29-DEC-1999.
 PD
 XX
 XX 25-JUN-1999; 99WO-NZ00092.
 PF
 XX
 XX 25-JUN-1998; 98US-0105307.
 PR
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 PA
 XX
 XX Havukkala JU, Bloksberg LN, Glenn M;
 PI
 XX
 XX WPI; 2000-116958/10.
 DR
 XX
 XX New plant microsatellite markers and associated flanking species for
 PT the detection of polymorphic genetic markers -
 PT
 XX
 XX Claim 1; Page 182; 392pp; English.
 PS
 XX
 XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences

CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms, in genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for
 CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.
 XX

SQ Sequence 344 BP; 72 A; 111 C; 93 G; 67 T; 1 other;
 Alignment Scores:
 Pred. No.: 6.05 Length: 344
 Score: 9.00 Matches: 9
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 1.81% Indels: 0
 DB: 21 Gaps: 0

US-09-872-523-5 (1-498) x AAA31397 (1-344)

QY 179 LysValGlnArgLeuGlnLeu 187
 |||||||
 Db 273 AAGTAGACACAGCTCGAATCTT 247

RESULT 9
 AAA31396/c
 ID AAA31396 standard; DNA; 350 BP.
 XX
 XX AAA31396;
 XX
 XX 05-JUL-2000 (first entry)
 DT
 XX
 XX Plant microsatellite marker #357.
 DE
 XX
 XX Plant microsatellite sequence; core repeat sequence: detection; probe;
 KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
 KW variety identification; genetic variability evaluation; primer; ss.
 XX
 XX Eucalyptus grandis.
 OS
 XX
 XX W09967421-A1.
 PN
 XX
 XX 29-DEC-1999.
 PD
 XX
 XX 25-JUN-1999; 99WO-NZ00092.
 PF
 XX
 XX 25-JUN-1998; 98US-0105307.
 PR
 XX
 XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
 PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
 PA
 XX
 XX Havukkala JU, Bloksberg LN, Glenn M;
 PI
 XX
 XX WPI; 2000-116958/10.
 DR
 XX
 XX New plant microsatellite markers and associated flanking species for
 PT the detection of polymorphic genetic markers -
 PT
 XX
 XX Claim 1; Page 182; 392pp; English.
 PS
 XX
 XX Sequences AAA31040-A32093 represent novel plant microsatellite sequences
 CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms, in genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for

CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.
 XX

SO Sequence 350 BP; 71 A; 113 C; 99 G; 67 T; 0 other:

Alignment Scores:

Pred. No.:	6.15	Length:	350
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	21	Gaps:	0

US-09-872-523-5 (1-498) x AAA31396 (1-350)

QY 179 LysValGlyGlnArgLeuGluLeu 187
 |||||
 Db 281 AAGGTAGACAGCGCTCGAACTTCTT 255

RESULT 10

AAA31294/C
 ID AAA31294 standard; DNA; 351 BP.

AC AAA31294;

DT 05-JUN-2000 (first entry)

XX Plant microsatellite marker #255.

DE Plant microsatellite sequence; core repeat sequence; detection; probe;

KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;

KW variety identification; genetic variability evaluation; primer; ss.

XX Euca1yptus grandis.

XX W09967421-A1.

XX 29-DEC-1999.

XX 25-JUN-1999; 99WO-N200092.

XX 25-JUN-1998; 98US-0105307.

XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.

XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Havukkala IJ, Bloksberg LN, Glenn M;

XX WPI: 2000-116958/10.

XX New plant microsatellite markers and associated flanking species for
 the detection of polymorphic genetic markers -

XX Claim 1; Page 151; 392pp; English.

CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
 CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms. In genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for
 CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.

SO Sequence 351 BP; 77 A; 111 C; 95 G; 68 T; 0 other:

Alignment Scores:

Pred. No.:	6.17	Length:	351
------------	------	---------	-----

Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	21	Gaps:	0

US-09-872-523-5 (1-498) x AAA31294 (1-351)

QY 179 LysValGlyGlnArgLeuGluLeu 187
 |||||
 Db 270 AAGGTAGACAGCGCTCGAACTTCTT 244

RESULT 11

AAA31381/C
 ID AAA31381 standard; DNA; 363 BP.

AC AAA31381;

DT 05-JUN-2000 (first entry)

XX Plant microsatellite marker #342.

DE Plant microsatellite sequence; core repeat sequence; detection; probe;

KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;

KW variety identification; genetic variability evaluation; primer; ss.

XX Euca1yptus grandis.

XX W09967421-A1.

XX 29-DEC-1999.

XX 25-JUN-1999; 99WO-N200092.

XX 25-JUN-1998; 98US-0105307.

XX (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.

XX (FLET-) FLETCHER CHALLENGE FORESTS LTD.

XX Havukkala IJ, Bloksberg LN, Glenn M;

XX WPI: 2000-116958/10.

XX New plant microsatellite markers and associated flanking species for
 the detection of polymorphic genetic markers -

XX Claim 1; Page 177; 392pp; English.

CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
 CC and associated flanking species. The sequences comprise a central core
 CC repeat sequence, especially selected from the sequences AAA32094-A32096
 CC with left and right flanking sequences. The polynucleotide sequences
 CC can be used in the detection of DNA polymorphisms. In genome mapping,
 CC in physical mapping, in positional cloning of genes, in variety
 CC identification and in evaluation of genetic variability within and
 CC between plant tissues, populations, cultivars, species and species
 CC groups. They may also be used to design hybridization probes for
 CC oligonucleotide fingerprinting and library screening and to design
 CC primers for microsatellite-primed PCR. Microsatellite markers are
 CC useful to locate specific economically useful genes in plant genomes.

SO Sequence 363 BP; 79 A; 114 C; 98 G; 72 T; 0 other:

Alignment Scores:

Pred. No.:	6.38	Length:	363
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	21	Gaps:	0

US-09-872-523-5 (1-498) x AAA31381 (1-363)

OY 179 LysVaIGlYGlAArgLeuGlLeuLeu 187
DB 276 AAGGTAGACAGCGCTCTCGAACTTCTT 250

RESULT 12
AAA31392/C
ID AAA31392 standard; DNA: 367 BP.
XX
AC AAA31392;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #353.
XX
KM Plant microsatellite sequence; core repeat sequence; detection; probe;
KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KM variety identification; genetic variability evaluation; primer; ss.
XX
OS Eucalyptus grandis.
XX
PN MO9967421-A1.
XX
PD 29-DEC-1999.
XX
PE 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 181; 392pp; English.
XX
CC Sequences AAA31040-A33093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SO Sequence 367 BP; 76 A; 116 C; 103 G; 72 T; 0 other;

Alignment Scores:
Pred. No.: 6.45 Length: 367
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 21 Gaps: 0

US-09-872-523-5 (1-498) x AAA31392 (1-367)

OY 179 LysVaIGlYGlAArgLeuGlLeuLeu 187
DB 256 AAGGTAGACAGCGCTCTCGAACTTCTT 230

RESULT 13
AAA31276/C
ID AAA31276 standard; DNA: 383 BP.
XX

AC AAA31276;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #237.
XX
KM Plant microsatellite sequence; core repeat sequence; detection; probe;
KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KM variety identification; genetic variability evaluation; primer; ss.
XX
OS Eucalyptus grandis.
XX
PN MO9967421-A1.
XX
PD 29-DEC-1999.
XX
PE 25-JUN-1999; 99WO-NZ00092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI; 2000-116958/10.
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
XX
PS Claim 1; Page 145; 392pp; English.
XX
CC Sequences AAA31040-A33093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SO Sequence 383 BP; 84 A; 120 C; 105 G; 74 T; 0 other;

Alignment Scores:
Pred. No.: 6.72 Length: 383
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 21 Gaps: 0

US-09-872-523-5 (1-498) x AAA31276 (1-383)

OY 179 LysVaIGlYGlAArgLeuGlLeuLeu 187
DB 270 AAGGTAGACAGCGCTCTCGAACTTCTT 244

RESULT 14
AAA31391/C
ID AAA31391 standard; DNA: 401 BP.
XX
AC AAA31391;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #352.
XX
KM Plant microsatellite sequence; core repeat sequence; detection; probe;
KM DNA polymorphism; genome mapping; physical mapping; fingerprinting;
XX

KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Eucalyptus grandis.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX
PF 25-JUN-1999; 99WO-N200092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
PI
DR WPI: 2000-116958/10.
XX
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
PS
PS Claim 1; Page 180; 392pp; English.
XX
CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 401 BP; 88 A; 124 C; 111 G; 78 T; 0 other;

Alignment Scores:
Pred. No.: 7.03 Length: 401
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 21 Gaps: 0

US-09-872-523-5 (1-498) x AAA31391 (1-401)
QY 179 LysValIGlyGlnArgLeuGluLeu 187
DB 264 AAGTAGACACGCTCTCGAACTTCTT 238

RESULT 15
AAA31390/C
ID AAA31390 standard: DNA; 407 BP.
XX
AC AAA31390;
XX
DT 05-JUL-2000 (first entry)
XX
DE Plant microsatellite marker #351.
XX
XX Plant microsatellite sequence; core repeat sequence; detection; probe;
KW DNA polymorphism; genome mapping; physical mapping; fingerprinting;
KW variety identification; genetic variability evaluation; primer; ss.
XX
OS Eucalyptus grandis.
XX
PN WO9967421-A1.
XX
PD 29-DEC-1999.
XX

PF 25-JUN-1999; 99WO-N200092.
XX
PR 25-JUN-1998; 98US-0105307.
XX
PN (GENE-) GENESIS RES & DEV CORP LTD & FLETCHER.
PA (FLET-) FLETCHER CHALLENGE FORESTS LTD.
XX
PI Havukkala IJ, Bloksberg LN, Glenn M;
XX
DR WPI: 2000-116958/10.
XX
XX
PT New plant microsatellite markers and associated flanking species for
PT the detection of polymorphic genetic markers -
PS
PS Claim 1; Page 180; 392pp; English.
XX
CC Sequences AAA31040-A32093 represent novel plant microsatellite sequences
CC and associated flanking species. The sequences comprise a central core
CC repeat sequence, especially selected from the sequences AAA32094-A32096
CC with left and right flanking sequences. The polynucleotide sequences
CC can be used in the detection of DNA polymorphisms, in genome mapping,
CC in physical mapping, in positional cloning of genes, in variety
CC identification and in evaluation of genetic variability within and
CC between plant tissues, populations, cultivars, species and species
CC groups. They may also be used to design hybridization probes for
CC oligonucleotide fingerprinting and library screening and to design
CC primers for microsatellite-primed PCR. Microsatellite markers are
CC useful to locate specific economically useful genes in plant genomes.
XX
SQ Sequence 407 BP; 87 A; 129 C; 113 G; 78 T; 0 other;

Alignment Scores:
Pred. No.: 7.14 Length: 407
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 21 Gaps: 0

US-09-872-523-5 (1-498) x AAA31390 (1-407)
QY 179 LysValIGlyGlnArgLeuGluLeu 187
DB 270 AAGTAGACACGCTCTCGAACTTCTT 244

Search completed: September 2, 2003, 20:47:27
Job time : 434 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 20:33:10 : Search time 94 Seconds
(without alignments)
2338.391 Million cell updates/sec

Title: US-09-872-523-5
Perfect score: 498
Sequence: 1 MSELKIVRANKSDRLDK.....PIGCEAHSYVLQPKKKNY 498

Scoring table:
OLIGO
Xgapop 60.0 , Ygapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 569978 seqs, 220691566 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1135299

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Command line parameters:

-MODL=frame+ p2n.model -DEV=xlh
-O=/cgn2_1/USPRO.spool/US09872523/runat_02092003_073003_27299/app.query.fasta.1.647
-DB=Issued_Patents_NA -OPMT=fastalg -SUFFIX=oli.rni -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL
-OUTPM=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09872523.ecgn.1.1.56.ernat.02092003_073003_27299 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESOURCY -NEG_SCORES=0 -NAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FCGAPOP=6
-FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

1: Issued_Patents_NA:*
2: /cgn2_6/ptodata/2/ina/5A.COMB.seq:*
3: /cgn2_6/ptodata/2/ina/5B.COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6A.COMB.seq:*
5: /cgn2_6/ptodata/2/ina/6B.COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	8	1.6	3526	1 US-08-323-084A-7	Sequence 7, Appl1
2	8	1.6	3526	1 US-08-323-084A-8	Sequence 8, Appl1
3	8	1.6	3526	1 US-08-674-008-7	Sequence 7, Appl1
4	8	1.6	3526	1 US-08-674-008-8	Sequence 8, Appl1
5	8	1.6	3986	4 US-09-053-374A-3	Sequence 3, Appl1
6	7	1.4	40	5 PCT-US94-06079-26	Sequence 26, Appl1
7	7	1.4	42	5 PCT-US94-06079-6	Sequence 6, Appl1
8	7	1.4	48	1 US-07-834-539A-64	Sequence 64, Appl1
9	7	1.4	48	1 US-08-053-131-112	Sequence 112, App
10	7	1.4	48	1 US-08-645-641-112	Sequence 112, App
11	7	1.4	48	1 US-07-853-408B-112	Sequence 112, App
12	7	1.4	48	1 US-08-096-762-112	Sequence 112, App

13	7	1.4	48	2 US-08-800-353-64	Sequence 64, Appl1
14	7	1.4	48	2 US-08-308-865-112	Sequence 112, App
15	7	1.4	48	3 US-09-042-353-305	Sequence 305, App
16	7	1.4	48	4 US-08-758-417A-153	Sequence 153, App
17	7	1.4	48	5 PCT-US92-06185-64	Sequence 64, Appl1
18	7	1.4	48	5 PCT-US92-10983-112	Sequence 112, App
19	7	1.4	119	1 US-09-465-901-5	Sequence 5, Appl1
20	7	1.4	119	1 US-07-634-278-80	Sequence 80, Appl1
21	7	1.4	119	1 US-08-477-728-80	Sequence 80, Appl1
22	7	1.4	119	1 US-08-474-040-80	Sequence 80, Appl1
23	7	1.4	119	1 US-08-487-200-80	Sequence 80, Appl1
24	7	1.4	119	3 US-08-484-537-80	Sequence 80, Appl1
25	7	1.4	146	1 US-08-116-388-9	Sequence 9, Appl1
26	7	1.4	146	5 PCT-US94-09318-9	Sequence 9, Appl1
27	7	1.4	162	4 US-08-746-411A-3	Sequence 3, Appl1
28	7	1.4	162	4 US-08-857-046A-3	Sequence 3, Appl1
29	7	1.4	162	4 US-09-573-252-3	Sequence 3, Appl1
30	7	1.4	248	5 PCT-US94-06079-8	Sequence 8, Appl1
31	7	1.4	273	4 US-09-313-294A-4219	Sequence 4219, App
32	7	1.4	275	4 US-09-313-294A-2176	Sequence 2176, App
33	7	1.4	282	4 US-09-216-393B-48	Sequence 48, Appl1
34	7	1.4	297	4 US-09-313-294A-2902	Sequence 2902, App
35	7	1.4	312	4 US-08-633-109-10	Sequence 10, Appl1
36	7	1.4	318	4 US-08-635-109-11	Sequence 11, Appl1
37	7	1.4	318	4 US-08-635-109-11	Sequence 11, Appl1
38	7	1.4	321	2 US-08-378-939-13	Sequence 13, Appl1
39	7	1.4	321	3 US-08-838-682-17	Sequence 17, Appl1
40	7	1.4	321	3 US-08-838-682-18	Sequence 18, Appl1
41	7	1.4	321	3 US-08-895-914-17	Sequence 17, Appl1
42	7	1.4	321	3 US-08-895-914-18	Sequence 18, Appl1
43	7	1.4	321	3 US-09-240-274-197	Sequence 197, App
44	7	1.4	321	3 US-09-357-710A-17	Sequence 17, Appl1
45	7	1.4	321	3 US-09-357-710A-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-08-323-084A-7
Sequence 7, Application US/08323084A
Patent No. 5563048
GENERAL INFORMATION:
APPLICANT: HONJO, TASHIRO
APPLICANT: SHIROZU, MICHIO
APPLICANT: TADA, HIDEAKI
TITLE OF INVENTION: No. 5563048el Polypeptides and DNAs encoding them
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEAK & SEAS
STREET: 2100 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20037-3202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323.084A
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 280505/1993
FILING DATE: 14-OCT-1993
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)293-7060
TELEFAX: (202)293-7860
TELEX: 6491103
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3526 base pairs

TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
US-08-323-084A-7

Alignment Scores:

Pred. No.:	128	Length:	3526
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	1	Gaps:	0

US-09-872-523-5 (1-498) x US-08-323-084A-7 (1-3526)

OY 352 GIuAspSerPheProIleHisIle 359
DB 1563 GAAGATCTTCCCAATTCACATC 1586

RESULT 2

US-08-323-084A-8
Sequence 8, Application US/08323084A

Patent No. 5563048
GENERAL INFORMATION:

APPLICANT: HONJO, TASUKU

APPLICANT: SHIROZU, MICHIO

APPLICANT: TADA, HIDEAKI

TITLE OF INVENTION: No. 5563048el Polypeptides and DNAs encoding them

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/323, 084A

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: JP 280505/1993

FILING DATE: 14-OCT-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ. ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 3526 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

CELL LINE: FLEB14

FEATURE:

NAME/KEY: CDS

LOCATION: 80..361

IDENTIFICATION METHOD: by similarity to some other pattern

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 80..142

IDENTIFICATION METHOD: by similarity with known sequence or

IDENTIFICATION METHOD: to an established consensus

NAME/KEY: mat_peptide
LOCATION: 143..358
IDENTIFICATION METHOD: by similarity with known sequence or
IDENTIFICATION METHOD: to an established consensus
US-08-323-084A-8

Alignment Scores:

Pred. No.:	128	Length:	3526
Score:	8.00	Matches:	8
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.61%	Indels:	0
DB:	1	Gaps:	0

US-09-872-523-5 (1-498) x US-08-323-084A-8 (1-3526)

OY 352 GIuAspSerPheProIleHisIle 359
DB 1563 GAAGATCTTCCCAATTCACATC 1586

RESULT 3

US-08-674-008-7
Sequence 7, Application US/08674008

Patent No. 5756084
GENERAL INFORMATION:

APPLICANT: HONJO, Tasuku

APPLICANT: SHIROZU, Michio

APPLICANT: TADA, Hideaki

TITLE OF INVENTION: HUMAN STROMAL DERIVED

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS

STREET: 2100 Pennsylvania Avenue, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20037-3202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/674, 008

CLASSIFICATION: 435

PRIORITY APPLICATION DATA:

APPLICATION NUMBER: JP 280505/1993

FILING DATE: 14-OCT-1993

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202)293-7060

TELEFAX: (202)293-7860

TELEX: 6491103

INFORMATION FOR SEQ. ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 3526 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA

US-08-674-008-7

FEATURE:

NAME/KEY: CDS

LOCATION: 80..361

IDENTIFICATION METHOD: by similarity to some other pattern

FEATURE:

NAME/KEY: sig_peptide

LOCATION: 80..142

IDENTIFICATION METHOD: by similarity with known sequence or

IDENTIFICATION METHOD: to an established consensus

IDENTIFICATION METHOD: to an established consensus

US-09-872-523-5 (1-498) x US-08-674-008-7 (1-3526)

OY 352 GluAspSerPheProIleHisIle 359

DB 1563 GAAGATTCCTTCCCAATTCACATC 1586

RESULT 4

US-08-674-008-8

; Sequence 8, Application US/08674008

; Patent No. 5736084

; GENERAL INFORMATION:

; APPLICANT: HONJO, Tasuku

; APPLICANT: SHIROZU, Michio

; APPLICANT: TADA, Hideaki

; TITLE OF INVENTION: HUMAN STROMAL DERIVED

; NUMBER OF SEQUENCES: 20

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: SUGHRUE, MION, ZINN, MACPHEAR & SEAS

; STREET: 2100 Pennsylvania Avenue, N.W.

; CITY: Washington

; STATE: D.C.

; COUNTRY: U.S.A.

; ZIP: 20037-3202

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; FILING DATE: 1-JUL-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/323,084

; FILING DATE: 14-OCT-1994

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: JP 280505/1993

; FILING DATE: 14-OCT-1993

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (202)293-7060

; TELEFAX: (202)293-7860

; TELEX: 6491103

; INFORMATION FOR SEQ ID NO: 8:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3526 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; CELL LINE: FLEB14

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 80..361

; IDENTIFICATION METHOD: by similarity to some other pattern

; FEATURE:

; NAME/KEY: sig_peptide

; LOCATION: 80..142

; IDENTIFICATION METHOD: by similarity with known sequence

; IDENTIFICATION METHOD:

; NAME/KEY: mat_peptide

; LOCATION: 143..358

; IDENTIFICATION METHOD: by similarity with known sequence

; IDENTIFICATION METHOD: or to an established consensus

US-08-674-008-8

Alignment Scores:

Pred. No.: 128

Score: 8.00

Length: 3526

Matches: 8

Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.61%
DB: 1
Gaps: 0

US-09-872-523-5 (1-498) x US-08-674-008-8 (1-3526)

OY 352 GluAspSerPheProIleHisIle 359

DB 1563 GAAGATTCCTTCCCAATTCACATC 1586

RESULT 5

US-09-053-374A-3/C

; Sequence 3, Application US/09053374A

; Patent No. 6462177

; GENERAL INFORMATION:

; APPLICANT: YEN, KWANG-MU

; TITLE OF INVENTION: MAMMALIAN BLOOD LOSS-INDUCED GENE, KD312

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: AMGEN INC.

; STREET: ONE AMGEN CENTER DRIVE

; CITY: THOUSAND OAKS

; STATE: CA

; COUNTRY: US

; ZIP: 91320

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: PatentIn Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/053,374A

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: COOK, ROBERT R.

; REGISTRATION NUMBER: 31,602

; REFERENCE/DOCKET NUMBER: A-514

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3986 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (genomic)

US-09-053-374A-3

Alignment Scores:

Pred. No.: 144

Score: 8.00

Percent Similarity: 100.00%

Best Local Similarity: 100.00%

Query Match: 1.61%

DB: 4

Gaps: 0

US-09-872-523-5 (1-498) x US-09-053-374A-3 (1-3986)

OY 426 ArgValGlyLeuArgLeuGluAla 433

DB 375 AGGCTCGGCTGAGCTGAGCGC 352

RESULT 6

PCT-US94-06079-26/C

; Sequence 26, Application PC/TUS9406079

; GENERAL INFORMATION:

; APPLICANT: American Cyanamid Company, .

; TITLE OF INVENTION: Gene Insertion by Direct Ligation In

; NUMBER OF SEQUENCES: 45

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: American Cyanamid Company

; STREET: One Cyanamid Plaza

CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06079
FILING DATE: 27-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31969-00\PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 26:
SEQUENCE CHARACTERISTICS:
LENGTH: 40 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-06079-26

Alignment Scores:
Pred. No.: 18.9 Length: 40
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
Gaps: 0

US-09-872-523-5 (1-498) x PCT-US94-06079-26 (1-40)

Qy 428 GtyleuArGLEuAlaAla 434
|||||
Db 21 GGCTGAGGCTGAGGCTGCT 1

RESULT 7
PCT-US94-06079-6/c
Sequence 6, Application PC/TUS9406079
GENERAL INFORMATION:
APPLICANT: American Cyanamid Company,
TITLE OF INVENTION: Gene Insertion by Direct Ligation In
VITRO
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: American Cyanamid Company
STREET: One Cyanamid Plaza
CITY: Wayne
STATE: New Jersey
COUNTRY: USA
ZIP: 07470-8426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/06079
FILING DATE: 27-MAY-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Gordon, Alan M.
REGISTRATION NUMBER: 30,637
REFERENCE/DOCKET NUMBER: 31969-00\PCT

TELECOMMUNICATION INFORMATION:
TELEPHONE: 201-831-3244
TELEFAX: 201-831-3305
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 42 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US94-06079-6

Alignment Scores:
Pred. No.: 19.8 Length: 42
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
Gaps: 0

US-09-872-523-5 (1-498) x PCT-US94-06079-6 (1-42)

Qy 428 GtyleuArGLEuAlaAla 434
|||||
Db 21 GGCTGAGGCTGAGGCTGCT 1

RESULT 8
US-07-834-539A-64
Sequence 64, Application US/07834539A
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
APPLICANT: Kay, Robert M.
TITLE OF INVENTION: Transgenic Non-Human Animals Capable of
Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Steuart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/834,539A
FILING DATE: 1992-02-05
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-5
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-543-9600
TELEFAX: 415-543-5043
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-07-834-539A-64

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-07-834-539A-64 (1-48)

OY 158 GYGLnThSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

RESULT 9

US-08-053-131-112

; Sequence 112, Application US/08053131

; Patent No. 5661016

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5661016-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 197

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Townsend and Townsend Hourie and Crew

; STREET: One Market Plaza, Steuart Tower, Suite 200

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94105

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/053,131

; FILING DATE: 26-Apr-1993

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/990,860

; FILING DATE: 16-DEC-1992

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/810,279

; FILING DATE: 17-DEC-1991

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/853,408

; FILING DATE: 18-MAR-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, William M.

; REGISTRATION NUMBER: 30,223

; REFERENCE/DOCKET NUMBER: 14643-9-3

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (primer)

; US-08-053-131-112

Alignment Scores:

Pred. No.: 22.6 Length: 48

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.41% Indels: 0

DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-053-131-112 (1-48)

OY 158 GYGLnThSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

DB 23 GGACAGACTCTCACCACCA 43

RESULT 10

US-08-645-641-112

; Sequence 112, Application US/08645641

; Patent No. 5719032

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5719032-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 150

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: William M. Smith

; STREET: Two Embarcadero Center, 8th Floor

; CITY: San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94111-3834

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/645,641

; FILING DATE: 20-MAY-1996

; CLASSIFICATION: 800

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/904,068

; FILING DATE: 23-JUN-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, William M.

; REGISTRATION NUMBER: 30,223

; REFERENCE/DOCKET NUMBER: 14643-000913

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-326-2400

; TELEFAX: 415-326-2422

; INFORMATION FOR SEQ ID NO: 112:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 48 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA (primer)

; US-08-645-641-112

Alignment Scores:

Pred. No.: 22.6 Length: 48

Score: 7.00 Matches: 7

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 1.41% Indels: 0

DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-645-641-112 (1-48)

OY 158 GYGLnThSerLeuSerPro 164

DB 23 GGACAGACTCTCACCACCA 43

RESULT 11

US-07-853-408B-112

; Sequence 112, Application US/07853408B

; Patent No. 5789650

; GENERAL INFORMATION:

; APPLICANT: Lonberg, Nils

; TITLE OF INVENTION: Transgenic No. 5789650-Human Animals for

; TITLE OF INVENTION: Producing Heterologous Antibodies

; NUMBER OF SEQUENCES: 150

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: William M. Smith

STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/853,408B
FILING DATE: 19920318
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (primer)
US-07-853-408B-112

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-07-853-408B-112 (1-48)
QY 158 GYGlnThrSerLeuSerPro 164
DB 23 GGACGAGCTCTCCTCACA 43

RESULT 12
US-08-096-762-112
Sequence 112, Application US/08096762
Patent No. 5814318
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
TITLE OF INVENTION: Transgenic No. 5814318-Human Animals for
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 210
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Kourie and Crew
STREET: One Market Plaza, Stewart Tower, Suite 200
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/096,762
FILING DATE: 22-JUL-1993
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/053,131
FILING DATE: 26-APR-1993

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/990,860
FILING DATE: 16-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/904,068
FILING DATE: 23-JUN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/853,408
FILING DATE: 18-MAR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/810,279
FILING DATE: 17-DEC-1991
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.
REGISTRATION NUMBER: 30,223
REFERENCE/DOCKET NUMBER: 14643-9-4
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-326-2400
TELEFAX: 415-326-2422
INFORMATION FOR SEQ ID NO: 112:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (primer)
US-08-096-762-112

Alignment Scores:
Pred. No.: 22.6 Length: 48
Score: 7.00 Matches: 7
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.41% Indels: 0
DB: 1 Gaps: 0

US-09-872-523-5 (1-498) x US-08-096-762-112 (1-48)
QY 158 GYGlnThrSerLeuSerPro 164
DB 23 GGACGAGCTCTCCTCACA 43

RESULT 13
US-08-800-353-64
Sequence 64, Application US/08800353
Patent No. 5874299
GENERAL INFORMATION:
APPLICANT: Lonberg, Nils
TITLE OF INVENTION: Transgenic No. 5874299-Human Animals Capable of
TITLE OF INVENTION: Producing Heterologous Antibodies
NUMBER OF SEQUENCES: 77
CORRESPONDENCE ADDRESS:
ADDRESSEE: William M. Smith
STREET: One Market Plaza, Stewart Tower, Suite 2000
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/800,353
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/834,539
FILING DATE: 1992-02-05
ATTORNEY/AGENT INFORMATION:
NAME: Smith, William M.

```

/
/   REGISTRATION NUMBER: 30,223
/   REFERENCE/DOCKET NUMBER: 14643-5
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: 415-543-9600
/   TELEFAX: 415-543-5043
/   INFORMATION FOR SEQ ID NO: 64:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 48 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/   MOLECULE TYPE: DNA (genomic)
/   US-08-800-353-64

Alignment Scores:
Pred. No.:      22.6      Length:      48
Score:          7.00      Matches:      7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.41%      Indels:      0
DB:             2         Gaps:         0

US-09-872-523-5 (1-498) x US-08-800-353-64 (1-48)

QY      158 GYGLnThrSerLeuSerPro 164
DB      23 GGACAGACTTCACCTCACCACCA 43

RESULT 14
US-08-308-865-112
/   Sequence 112, Application US/08308865
/   Patent No. 5877397
/   GENERAL INFORMATION:
/   APPLICANT: Lonberg, Nils
/   TITLE OF INVENTION: Transgenic No. 5877397-Human Animals for
/   TITLE OF INVENTION: Producing Heterologous Antibodies
/   NUMBER OF SEQUENCES: 150
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: William M. Smith
/   STREET: One Market Plaza, Steuart Tower, Suite 2000
/   CITY: San Francisco
/   STATE: California
/   COUNTRY: USA
/   ZIP: 94105
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: Floppy disk
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: PatentIn Release #1.0, Version #1.25
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/308,865
/   FILING DATE:
/   CLASSIFICATION: 800
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US/08/145,707
/   FILING DATE:
/   APPLICATION NUMBER: US 07/904,068
/   FILING DATE: 23-JUN-1992
/   ATTORNEY/AGENT INFORMATION:
/   NAME: Smith, William M.
/   REGISTRATION NUMBER: 30,223
/   REFERENCE/DOCKET NUMBER: 14643-9-1-1
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: 415-326-2400
/   TELEFAX: 415-326-2422
/   INFORMATION FOR SEQ ID NO: 112:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 48 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/   MOLECULE TYPE: DNA (primer)
/

/
/   REGISTRATION NUMBER: 30,223
/   REFERENCE/DOCKET NUMBER: 14643-5
/   TELECOMMUNICATION INFORMATION:
/   TELEPHONE: 415-543-9600
/   TELEFAX: 415-543-5043
/   INFORMATION FOR SEQ ID NO: 64:
/   SEQUENCE CHARACTERISTICS:
/   LENGTH: 48 base pairs
/   TYPE: nucleic acid
/   STRANDEDNESS: single
/   TOPOLOGY: linear
/   MOLECULE TYPE: DNA (genomic)
/   US-08-800-353-64

Alignment Scores:
Pred. No.:      22.6      Length:      48
Score:          7.00      Matches:      7
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:    1.41%      Indels:      0
DB:             2         Gaps:         0

US-09-872-523-5 (1-498) x US-08-308-865-112 (1-48)

QY      158 GYGLnThrSerLeuSerPro 164
DB      23 GGACAGACTTCACCTCACCACCA 43

RESULT 15
US-09-042-353-305
/   Sequence 305, Application US/09042353
/   Patent No. 6255458
/   GENERAL INFORMATION:
/   APPLICANT: Lonberg, Nils
/   TITLE OF INVENTION: Transgenic No. 6255458-Human Animals for
/   TITLE OF INVENTION: Producing Heterologous Antibodies
/   NUMBER OF SEQUENCES: 421
/   CORRESPONDENCE ADDRESS:
/   ADDRESSEE: Townsend and Townsend and Crew LLP
/   STREET: Two Embarcadero Center, Eighth Floor
/   CITY: San Francisco
/   STATE: California
/   COUNTRY: USA
/   ZIP: 94111-3834
/   COMPUTER READABLE FORM:
/   MEDIUM TYPE: Floppy disk
/   OPERATING SYSTEM: PC-DOS/MS-DOS
/   SOFTWARE: PatentIn Release #1.0, Version #1.30
/   CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/09/042,353
/   FILING DATE: 13-MAR-1998
/   CLASSIFICATION: 800
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 07/810,279
/   FILING DATE: 17-DEC-1991
/   APPLICATION NUMBER: US 07/853,408
/   FILING DATE: 18-MAR-1992
/   APPLICATION NUMBER: US 07/904,068
/   FILING DATE: 23-JUN-1992
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 07/990,860
/   FILING DATE: 16-DEC-1992
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 08/053,131
/   FILING DATE: 26-APR-1993
/   APPLICATION NUMBER: US 08/096,762
/   FILING DATE: 22-JUL-1993
/   PRIOR APPLICATION DATA:
/   APPLICATION NUMBER: US 08/155,301
/   FILING DATE: 18-NOV-1993
/   APPLICATION NUMBER: US 08/161,739
/   FILING DATE: 03-DEC-1993
/   APPLICATION NUMBER: US 08/165,699
/   FILING DATE: 10-DEC-1993
/   APPLICATION NUMBER: US 08/209,741
/   FILING DATE: 09-MAR-1994
/   PRIOR APPLICATION DATA:
/
```

APPLICATION NUMBER: US 08/352,322
FILING DATE: 07-DEC-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/544,404
FILING DATE: 10-OCT-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/728,463
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US96/16433
FILING DATE: 10-OCT-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/758,417
FILING DATE: 02-DEC-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: WO PCT/US97/21803
FILING DATE: 01-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Apple, Randolph T.
REGISTRATION NUMBER: 36,429
REFERENCE/DOCKET NUMBER: 014643-009040US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 305:
SEQUENCE CHARACTERISTICS:
LENGTH: 48 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
US-09-042-353-305

Alignment Scores:

Pred. No.:	22.6	Length:	48
Score:	7.00	Matches:	7
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.41%	Indels:	0
DB:	3	Gaps:	0

US-09-872-523-5 (1-498) x US-09-042-353-305 (1-48)

QY 158 GlycInPrSerIeuSerPro 164
DB 23 GGACGACTCTCCTCTCACCA 43

Search completed: September 2, 2003, 23:12:18
Job time : 107 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 20:40:10 ; Search time 369 seconds
(without alignments)
3096.306 Million cell updates/sec

Title: US-09-872-523-5
Perfect score: 498
Sequence: 1 MSEFLKIVRANKSRDKLKD.....PIGWCEAHSYVLPPKKNY 498

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1533700 segs, 1147125425 residues

Word size: 1
Total number of hits satisfying chosen parameters: 3063161

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODE=frame+p2n.model -DEV=xlh
-O=/cgn2_1/USPTO.spool/US09872523/runat_02092003_073004_27329/app.query.fasta.1.647
-DB=Published.Applications_NA -QFMT=fastcap -SUFFIX=oli.rnpb -MINMATCH=0.1
-LOOPEC=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo
-TRANS=human4.0.coi -List=45 -DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1
-ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09872523.@CGN.1.1.244 @runat_02092003_073004_27329
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60
-FCGAPOP=6 -FCGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Published.Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	498	100.0	1497	10	US-09-872-523-6 Sequence 6, Appl1

2	411	82.5	1497	10	US-09-872-523-73	Sequence 73, Appl1
3	411	82.5	1497	10	US-09-872-523-74	Sequence 74, Appl1
4	397	79.7	1497	10	US-09-872-523-75	Sequence 75, Appl1
5	278	55.8	2307	10	US-09-872-523-76	Sequence 76, Appl1
6	278	55.8	2307	10	US-09-872-523-77	Sequence 77, Appl1
7	278	55.8	2307	10	US-09-872-523-78	Sequence 78, Appl1
8	234	1.6	234	10	US-09-878-574-13281	Sequence 13281, A
9	234	1.6	437	13	US-10-027-632-19030	Sequence 19030, A
10	590	1.6	590	13	US-10-027-632-226880	Sequence 226880, A
11	590	1.6	590	13	US-10-027-632-226881	Sequence 226881, A
12	608	1.6	608	13	US-10-027-632-222125	Sequence 222125, A
13	609	1.6	609	13	US-10-027-632-6153	Sequence 6153, Ap
14	611	1.6	611	13	US-10-027-632-224608	Sequence 224608, A
15	611	1.6	611	13	US-10-027-632-224609	Sequence 224609, A
16	632	1.6	632	13	US-10-027-632-141937	Sequence 141937, A
17	669	1.6	669	13	US-10-027-632-96511	Sequence 96511, A
18	669	1.6	669	13	US-10-027-632-306237	Sequence 306237, A
19	833	1.6	833	13	US-10-027-632-306237	Sequence 165347, A
20	833	1.6	833	13	US-10-144-929-63	Sequence 63, Appl1
21	3524	1.6	3524	8	US-08-927-939-29	Sequence 29, Appl1
22	3524	1.6	3524	10	US-09-919-497-45	Sequence 45, Appl1
23	3524	1.6	3524	10	US-09-880-107-2204	Sequence 2204, Ap
24	4313	1.6	4313	10	US-09-070-927A-67	Sequence 67, Appl1
25	9424	1.6	9424	8	US-08-910-386A-8	Sequence 8, Appl1
26	19011	1.6	19011	11	US-09-764-891-5813	Sequence 5813, Ap
27	19011	1.6	19011	11	US-09-764-891-5814	Sequence 5814, Ap
28	271990	1.6	271990	14	US-10-195-144-87	Sequence 87, Appl1
29	1601042	1.6	1601042	13	US-10-027-632-55064	Sequence 55064, A
30	87	1.4	87	14	US-10-259-430-5	Sequence 5, Appl1
31	201	1.4	201	13	US-10-027-632-24317	Sequence 24317, A
32	210	1.4	210	9	US-09-764-869-313	Sequence 313, Ap
33	210	1.4	210	14	US-10-091-504-313	Sequence 313, Ap
34	258	1.4	258	9	US-09-294-093B-781	Sequence 781, Ap
35	253	1.4	253	9	US-09-294-093B-3099	Sequence 3099, Ap
36	265	1.4	265	10	US-09-878-574-5748	Sequence 5748, Ap
37	269	1.4	269	10	US-09-796-692-3156	Sequence 3156, Ap
38	269	1.4	269	14	US-10-040-862-3156	Sequence 3156, Ap
39	275	1.4	275	9	US-09-923-876-5298	Sequence 5298, Ap
40	280	1.4	280	9	US-09-864-761-31987	Sequence 31987, A
41	282	1.4	282	9	US-09-878-574-15142	Sequence 15142, A
42	282	1.4	282	9	US-09-216-393-48	Sequence 48, Appl1
43	285	1.4	285	10	US-09-878-574-5400	Sequence 5400, Ap
44	285	1.4	285	10	US-09-878-574-12434	Sequence 12434, A
45	285	1.4	285	14	US-10-102-524-1259	Sequence 1259, Ap

ALIGNMENTS

RESULT 1
US-09-872-523-6
; Sequence 6, Application US/09872523
; Patient No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-6
Alignment Scores: 0
Pred. No.: 498.00
Length: 1497
Matches: 498

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-6 (1-1497)

```
OY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
Db 1 ATGCTGAATTTTCGAAAAATGTCAGAGCTTAACAAAAATCGGAGAAAACTCGATTAAG 60
OY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
Db 61 ACCTACTGTGGGAATCCTATTACATCAGTTCCGAAAGGAAAAAATCTTTTCATTTCCA 120
OY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValLysGluIleVal 60
Db 121 GTTGAGAGATTCATTCGTAACCTTAACGTAATTTTAACGAATGGCTGAGAGAGAGATT 180
OY 61 IlePheGluThrValIleHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyr 80
Db 181 ATCTTCGAAACAGAGTGCATGATTAATGACAAAGAACTCGATTCAAGTCAGATG 240
OY 81 PheIleArgIleGluLysValCysGlyTyrArgValIleuAlaGlnPheIleGlyAlaAsp 100
Db 241 TTTCGACGAATTGAAAAAGTTTGGGATACAGAGTTCTGGCTCAGATTATGGAGCTGAC 300
OY 101 ThrLysPheTyrPheLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
Db 301 ACGAAATTTTGGCCCAATATTTTATCGACAGCATATGTTGGTTGGCAAGCCCGCAATG 360
OY 121 SerAspProAsnMetAspLysIleValTyrAlaProPheLeuAlaIleAsnGluLysTyr 140
Db 361 AGTAGTCCCAATATGGATGAAATGTAATGTCCTCGCCGCTTGCAATACGAAAGAAATAC 420
OY 141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
Db 421 CAAATATGATATGTAATTAATGTAATTAATGTCATGATGGCGAAATGCTGCGCCAACT 480
OY 161 SerLeuSerProLysPheAspGluLysAlaLeuLeuSerLysHisArgPheLysVal 180
Db 481 TCGGTGCTCCAAATTCGATGAAAGGAGGCTCTCTAAGCAACGATCGTTTCAAGCTT 540
OY 181 GlyGlnArgLeuGlnLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
Db 541 GGACCAACCTCTTGAACATTAATTAATTCATTCCTGTAAGTAACGGGTAGCCGGAAT 600
OY 201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
Db 601 CAAAGAAATATGTGACGACGACGATGATATCTATCAAAAGAAAGACTTTCCCAATCG 660
OY 221 LeuProAspAlaAspAspAspArgIleValIlePheSerSerGlySerGluTyrTrpIleAsp 240
Db 661 CTTCGACAGTACGATGACGACGACGACAAAGCTTTAGCTCTGATCTCAATATGATGAT 720
OY 241 GlnGlySerPhePheIlePheProValGlyPheAlaIleValAsnGlyTyrGlnLeuAsn 260
Db 721 GAGGGAACTCTTCATATTCCTCTGTTGATTTGGAGAGTCGAAATGATCACTAAAT 780
OY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
Db 781 GCGAAAAAGGAATATATGAGACACCAATTAATTTGCTCAACGAATAAAAATGAGAGA 840
OY 281 AsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAsp 300
Db 841 AATCCAAATATGACTCAGACGACCTCAATTTGATCAATAGCAAAAGATCCCAATGAT 900
OY 301 ProMetIleTrpArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
Db 901 CCCATGATTTGGAGAAAGTTAAGTTGAGACAAAAGTTTGAAGCTATGACCCCTTGCT 960
OY 321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluLysTyr 340
```

```
Db 961 CAGCAATTCATTAACCTCCAGTCGCTTCGATTTCGAAATTTTGCAAAACCTGAGAGATAT 1020
OY 341 LeuIleValGlyMetAspGlyProAspAlaLeuGlnAspSerPheProIleHisIleAsn 360
Db 1021 CTTAATTTGGGAGATGAGATGTCAGATGCTGACCTTGAAGACAGTTTCCATTAATCAATAT 1080
OY 361 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGlnLeuValProPro 380
Db 1081 AATCAATTAATGTTCCCGAGTTGGTTATCGGAAAAAGTAAATTTTGGAACTGTTCCCGCA 1140
OY 381 AspGluPheLysGlyThrPheArgTyrAspGluTyrLeuGluLysGluSerAlaGluThr 400
Db 1141 GATGACTTCGAAGACAACTTATGATGAGATGAATTAATCTTGAGAAAAGATTCGCAAAAC 1200
OY 401 LeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVal 420
Db 1201 CTACCGCTTGACTGTTTCAACGCAATGCTTCCCAAGAGAGATTAGCAAAATTTAAGSTA 1260
OY 421 IleLeuIleSerLysArgValGlyLeuArgLeuGluAlaIleAspMetCysGluAsnGln 440
Db 1261 ATTCGATTTCCAAACGGGTAGGACTACGCTTGAAAGCTGCTGACATGTGAAAAATCAG 1320
OY 441 PheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAsp 460
Db 1321 TTTATTTTCCAGCTACAGTGAATCAATCACTTATGAGAACTGATTAATCTCAATTTGAC 1380
OY 461 GlyTyrAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProIle 480
Db 1381 GGCTGGAGTGAAGAAATTTGATGAACCTGATGATGGAGCTCCCATGATATTTCCACGATA 1440
OY 481 GlyTyrCysGluAlaHisSerTyrValLeuGlnProPheLysLysTyrAsnTyr 498
Db 1441 GGATGGTGTGAAGCGCACAGTATGTTCTACAACTCCGAAAAAGATACAACTAT 1494

RESULT 2
US-09-872-523-73
; Sequence 73, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davidson, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 73
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-73

Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 411.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 82.53% Indels: 2
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-73 (1-1497)
OY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
Db 1 ATGCTGAATTTTCGAAAAATGTCAGAGCTTAACAAAAATCGGAGAAAACTCGATTAAG 60
OY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
Db 61 ACCTACTGTGGGAATCCTATTACATCAGTTCCGAAAGGAAAAAATCTTTTCATTTCCA 120
```

QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValIysGluGlyVal 60
 |||||
 Db 121 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAAGGAGAGAGATT 180
 QY 61 IlePheGluThrValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyr 80
 |||||
 Db 181 ATCTTCGAAACAGTGGTCCATGATTTATGACAAAGAACTGGCATTCGATTCAGATGCG 240
 QY 81 PheAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAsp 100
 |||||
 Db 241 TTTCGACGAAATGAAAGATTTCGGATACAGAGTCTCGCTCAGTTTATCGGAGCTGAC 300
 QY 101 ThrLysPheTyrLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120
 |||||
 Db 301 ACGAAATTTTGGCTCATATTTTATCGGACGATATCTTTGGTTGGCAAAACCGCCAAATG 360
 QY 121 SerAspProAsnMetAspLysIleValTyrAlaProPheLeuAlaIleAsnGluTyr 140
 |||||
 Db 361 AGTGATCCCAATATGATGATAAATTGTATATGCTCCGCCGTTGCATTCACAGAAATAC 420
 QY 141 GlnAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
 |||||
 Db 421 CAAAAGATATGTAATAATTATGTAATAATTGCATTGATGCCAAATTCGCGCCAAAGT 480
 QY 161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
 |||||
 Db 481 TCGCTCTCCAAAATTCATGAGGAAAGCGCTCTCTTAAGCAAGCATGCTTTCAAGTT 540
 QY 181 GlyGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
 |||||
 Db 541 GGACAAACGCTTGAACATATAATTAATTAATTCCTCAATTCAGTGAATACGCCGAGAAAT 600
 QY 201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
 |||||
 Db 601 CAAAGAAATATGTGGACGAGCAAGAAATGATATCTATCACAAAGAAACCTTTCCCGAATG 660
 QY 221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAsp 240
 |||||
 Db 661 CTTCCAGATGCAAGATACGACAGCAAGCTTTAGCTCGAGATCTCAATATTTGATGATGAC 720
 QY 241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
 |||||
 Db 721 GAGGAGAGCTTCTTCAATATTTCTTGGATTGTCAGCGATTCATGATCAACTAAT 780
 QY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyGlu 280
 |||||
 Db 781 GCGAAAGAAAGAAATATATTTAGCACACAAATTAATTTGCTCAACCAATTAATAATGAGAA 840
 QY 281 AsnProArgTyrAspSerAspAspValThrPheAspGluLeuAlaLysAspProIleAsp 300
 |||||
 Db 841 AATCCAAAGATATGACATCACACAGCATGATTTGATTCATTAACAAAGATTCATTTGAT 900
 QY 301 PrometIleTyrArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
 |||||
 Db 901 CCCATGATTTGGGAAAGAAATTAAGTTGAAGAAAGTTTACGCTCATACACCCCTTGCT 960
 QY 321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGlyTyr 340
 |||||
 Db 961 CAGCAATTCATTAACCTCCACAGCTCGCTTCGATTTCAATTTTGGCAAACTGAAAGATAT 1020
 QY 341 IleuIleValGlyMetAspGlyProAspAlaLeuGluLysSerPheProIleHisIleAsn 360
 |||||
 Db 1021 CTTATTTGGGAATGATGCTCAGATGACACTTGAAGACAGTTTCTTCTATCATCAT 1080
 QY 361 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPro 380
 |||||
 Db 1081 AATACATTTATGTCTCCAGTTGGTTATGCGGAAAGAAATTAATTTGGAACCTTGTCCGCA 1140
 QY 381 AspGluPheLysGlyThrPheArgTyrPaspGluTyrLeuGluLysGluSerAlaGluThr 400
 |||||
 Db 1141 GATGAGTTCAAGGAAACATTCAGATGGGATGATTACTTGGAGAAAGAAATCTGCAGAAACC 1200

QY 401 LeuProLeuAspLeuPheLysPrometProSerGln-GluArgLeuAspLysPheLysVal 420
 |||||
 Db 1201 CTACCGCTTGACTGTTCAAGCCAAATGCCCTTCTTA-AGACAGATTTGACAAATTTAAGG 1259
 QY 420 IleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnGly 440
 |||||
 Db 1260 AATTCGATTTCCAAACGGGTAGACTACGCTTGAAGCGTCGACATGTGTGAAATCA 1319
 QY 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
 |||||
 Db 1320 GTTTATTTGCGCAGCTACAGTGAATCAGTTTCATGGAAGACTGATTAATGTCATTTTCCA 1379
 QY 460 pGlyTyrPaspGluGlnPheAspGluLeuTyrTyrAspValAspSerHisAspIleLeuProI 480
 |||||
 Db 1380 CGGCTGGATGGAAGAAATTTGATGACTGATGATGATGATGATGATGATGATGATGATGAT 1439
 QY 480 eGlyTyrCysGluAlaHisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
 |||||
 Db 1440 AGGATGGTGGAAGCGCACAGTTATGTTCTACAACTCCGAAAGATCAACTAT 1494

RESULT 3

US-09-872-523-74
 ; Sequence 74, Application US/09872523
 ; Patent No. US20020137906A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Horvitz, H. Robert
 ; APPLICANT: Davison, Ewa M.
 ; APPLICANT: Lu, Xiaowei
 ; TITLE OF INVENTION: A tumor suppressor pathway in C. Elegans
 ; FILE REFERENCE: 0197/536002
 ; CURRENT APPLICATION NUMBER: US/09/872,523
 ; PRIORITY FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: US 60/208,802
 ; NUMBER OF SEQ ID NOS: 78
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 74
 ; LENGTH: 1497
 ; TYPE: DNA
 ; ORGANISM: Caenorhabditis elegans
 US-09-872-523-74

Alignment Scores:
 Pred. No.: 0 Length: 1497
 Score: 411.00 Matches: 497
 Percent Similarity: 99.60% Conservative: 0
 Best Local Similarity: 99.60% Mismatches: 1
 Query Match: 82.53% Indels: 2
 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-74 (1-1497)

QY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
 |||||
 Db 1 ATGTCTGAATTTCTGAAATTTCTCAGAGCTTACAAAGAAATTCGACAGAAATTCGATTAAG 60
 QY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
 |||||
 Db 61 ACCTTACTTGGGAAATCTTATTTACATTCATTCATTCGAGAAAGAAAGAAACTTTTTCATTCCA 120
 QY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPheAsnGluCysValIysGluGlyVal 60
 |||||
 Db 121 GTTGAAGCATTCATCGTAACCTTACAGTTAATTTTAAAGAAAGGAGAGATT 180
 QY 61 IlePheGluThrValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTyr 80
 |||||
 Db 181 ATCTTCGAAACAGTGGTCCATGATTTATGACAAAGAACTGCGATTCGATTCAGATGCG 240
 QY 81 PheAlaArgIleGluLysValCysGlyTyrArgValLeuAlaGlnPheIleGlyAlaAsp 100
 |||||
 Db 241 TTTCGACGAAATTAAGAAAGTTTCGGATACAGAGTTCTGCTAGTTATCGGAGCTGAC 300
 QY 101 ThrLysPheTyrLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaIleMet 120

|||||
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACATATGTTGGTGGCAACGCCCAATG 360
OY SeraspproAsnMetAspLysIleValTyrAlaProProLeuAlaIleAsnGluGluTyr 140
Db 361 AGTATGCCAATATGATGAATAATGCTATATGCTCCGCCGCTTGCAATACAGAAAGATAC 420
OY 141 GluAsnAspMetValAsnTyrValAsnAsnCysIleAspGluGluIleValGlyGlnThr 160
Db 421 CAAATGATATGTGTAATTTATGTAATGATTCATTCATTCAGGAATGTCGGCCAAACT 480
OY 161 SerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
Db 481 TCGCTGCTCCAAATTCGATGAGAGGAGGCTCTCCCTAGCAACCATCGTTTCAAAGT 540
OY 181 GlyGlnArgLeuGluLeuLeuAsnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
Db 541 GGACACAGCTTGAACCTATTAATATTCATTCATTCATTCATTCATTCATTCATTCATTC 600
OY 201 GlnGluIleCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSer 220
Db 601 CACGAAATATGTGCGACGACGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
OY 221 LeuProAspAlaAspAspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAsp 240
Db 661 CTTCAGATGCGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 720
OY 241 GluGlySerPhePheIlePheProValGlyPheAlaAlaValAsnGlyTyrGlnLeuAsn 260
Db 721 GAGGAGACCTTCTTCATATTCCTGATGATGATGATGATGATGATGATGATGATGATGAT 780
OY 261 AlaLysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlu 280
Db 781 GCGAAAAAGGATATATGATGACACCAATATAATGCTCAAGCAATATAATATAATATAAT 840
OY 281 AsnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAsp 300
Db 841 AATCCAAATATGACTCGACGACGCTCATTGATTCATTCATTCATTCATTCATTCATTCAT 900
OY 301 ProMetIleTrrPargLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAla 320
Db 901 CCCATGATTTGGAGAAAGATTAAAGCTTGACAAAAGTTGAGCTCATGACCCCTTGCT 960
OY 321 GlnGlnPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGlnGluTyr 340
Db 961 CAGCAATTCATTAACCTCCAGCTGCTTCGATTCCTCAATTTTGCAAAACCTGAAGAGAT 1020
OY 341 LeuIleValGlyMetAspGlyProAspAlaLeuGluAspSer-PheProIleHisIleAs 360
Db 1021 CTTATGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1079
OY 360 AsnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValProPr 380
Db 1080 TAAATCAATTTATGTTCCAGTGGTTATGCGGAAAAGTATAATTTGGAACCTTGTCGCC 1139
OY 380 AspGluPheLysGlyTrrPheArgTrrPaspGluTyrLeuGluLysGluSerAlaGluThr 400
Db 1140 AGATAGATTCAAGGAAACATTCAGATGGATGATGATGATGATGATGATGATGATGATGAT 1199
OY 400 rLeuProLeuAspLeuPheLysProMetProSerGlnGluArgLeuAspLysPheLysVa 420
Db 1200 CCTACCGCTTGACTTGTCACAGCCATGCTTCCCAAGAGAGATTAAGCAATTTTAAGGT 1259
OY 420 IleLeuIleSerLysArgValGlyLeuArgLeuGluAlaAlaAspMetCysGluAsnGlu 440
Db 1260 AATTCGATTTCCAAACGGGTAGGACTAGCCTTGAGCTGTCGATCTGTAAGTGAATCA 1319
OY 440 nPheIleCysProAlaThrValLysSerValHisLysArgLeuIleAsnValAsnPhes 460
Db 1320 GTTATATTTGCTCCAGCTACAGTAAATCAGTTTCATGGAAGACATGATTAATTAATTCGA 1379
OY 460 PGIYTrpAspGluGluPheAspGluLeuTyrAspValAspSerHisAspIleLeuProIle 480

Db 1380 CGCGTGGGATGAGCAATTTGATGAACCTGATGATGTGGACTCCCATGATATTCTACCGAT 1439
OY 480 eGIYTrpCysGluAlaHisSerTyrValLeuGlnProProLysTyrAsnTyr 498
Db 1440 AGGATGGTGTGAAGGCGACAGTATGTTCTTCACAACTCCGAAAAAGTACAACTAT 1494
RESULT 4
US-09-872-523-75
; Sequence 75, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Eva M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75
; LENGTH: 1497
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-75
Alignment Scores:
Pred. No.: 0 Length: 1497
Score: 397.00 Matches: 497
Percent Similarity: 99.60% Conservative: 0
Best Local Similarity: 99.60% Mismatches: 1
Query Match: 79.72% Indels: 2
DB: 10 Gaps: 0
US-09-872-523-5 (1-498) x US-09-872-523-75 (1-1497)
OY 1 MetSerGluPheLeuLysIleValArgAlaAsnLysLysSerAspArgLysLeuAspLys 20
Db 1 ATGTGTGAATTTCTGAAAAATGTCAGAGCTTAACAAAAATCGACAGAAACTCGATAG 60
OY 21 ThrTyrLeuTrpGluSerTyrLeuHisGlnPheGluLysGlyLysThrSerPheIlePro 40
Db 61 ACCTACTGTGGGAATCTCATTTACATCAGTTCCGAGAAAGGAAAACTCTTTCATTCCA 120
OY 41 ValGluAlaPheAsnArgAsnLeuThrValAsnPhesnGluCysValLysGluGlyVal 60
Db 121 GTTGAAGCATTCATCGTAACTTACCTTACAGTTAATTTTAACGAATCGGTGAAGAGAGT 180
OY 61 IlePheGluThrValValHisAspTyrAspLysAsnCysAspSerIleGlnValArgTrp 80
Db 181 ATCTTCGAAACGTTGCTCATGATTAATGACAAGAACTCGATTCATTCAGTCAGATGG 240
OY 81 PheAlaArgIleGluLysValCysGlyTyrArgValIleAlaGlnPheIleGlyAlaAsp 100
Db 241 TTTGACACGATTTGAAAAAGTTTGGCGATACAGAGTTGCGTCATATTATGAGACTGAC 300
OY 101 ThrLysPheTrpLeuAsnIleLeuSerAspAspMetPheGlyLeuAlaAsnAlaMet 120
Db 301 ACGAAATTTTGGCTCAATATTTTATCGGACGATGATGTTGGTGGCAAAAGCCCAATG 360
OY 121 SerAspProAsnMetAspLysIleValTyrAlaPro-ProLeuAlaIleAsnGluGluTyr 140
Db 361 AGTGATCCCAATATGATTAATTTATATATGCTTC-GCCGCTTGCAATCAACGAAGAATA 419
OY 140 rGluAsnAspMetValAsnTyrValAsnAsnCysIleAspGlyGluIleValGlyGlnThr 160
Db 420 CCAAAATGATATGTAATTTATGTAATATATTCATTCATTCATTCATTCATTCATTCATTC 479
OY 160 rSerLeuSerProLysPheAspGluGlyLysAlaLeuLeuSerLysHisArgPheLysVal 180
Db 480 TTCGCTGCTCCAAATTCGATGAGGAGGAGGCTCTCCTAAGCAAGCATCGTTCAAAGT 539

QY 180 lglglnargleuGluleuLeuasnTyrSerAsnSerThrGluIleArgValAlaArgIle 200
Db 540 TGGACAACTCTTGAACCTTAATTAATTCCTCAATTCCTGAAATACGGGTACGGGAAT 599
QY 200 eglngluileCysGlyArgArgMetAsnValSerIleThrLysLysAspPheProGluSe 220
Db 600 TCAGAAATATGTGACGACGAATGAATGATCTATCACAAAGAAAGACTTCCCGAATC 659
QY 220 rleupProaspAlaaspAspArgGlnValPheSerSerGlySerGlnTyrTrpIleAs 240
Db 660 GCTTCAGATGAGATGACGACGACAGACAACTTCTTACCTGTGGATCTCAATATTGGATAGA 719
QY 240 psluglSerPhePheIlePheProValGlyPheAlaValAsnGlyTyrGlnLeuAs 260
Db 720 CGAGGAACCTTCTTCATATTTCCTGTGATTCGACGACGATCAATGATACACTAA 779
QY 260 naLalysLysGluTyrIleGluHisThrAsnLysIleAlaGlnAlaIleLysAsnGlyG 280
Db 780 TCGGAAAAAGAAATATATTGAGCACACAAATAAATTTGCTCAAGCAATTAATAATGAGA 839
QY 280 uasnProArgTyrAspSerAspAspValThrPheAspGlnLeuAlaLysAspProIleAs 300
Db 840 AATCCAAATATGACTCAGACGACGCTCACATTTGATTAAGCAAAAGATCCAAATGA 899
QY 300 pPromeIleTrpArgLysValLysValGlyGlnLysPheGluLeuIleAspProLeuAl 320
Db 900 TCCCATGATTTGGAGAAAAGTTAAGTTGACAAAAGTTGAGCTCATGACCCCTTGGC 959
QY 320 actInglInPheAsnAsnLeuHisValAlaSerIleLeuLysPheCysLysThrGluGly 340
Db 960 TCAGCAATTCATTAACCTCCAGCTCCCTCGATTCCTCAATTTTGCAGAAAAGTAAAGANA 1019
QY 340 rleuIleValGlyMetAspGlyProaspAlaLeuGluAspSerPheProIleHisLys 360
Db 1020 TCTTATTGGGAGATGATGCTCAGATCAGCTTGAAGACAGTTTCTCATTCATATCA 1079
QY 360 nasnThrPheMetPheProValGlyTyrAlaGluLysTyrAsnLeuGluLeuValIlePro 380
Db 1080 TAAATCATTTATGTTCCAGTTGGTTATCGGAAAAGTTAATTTGCACTTGTCCGCC 1139
QY 380 oaspGluPheLysGlyThrPheArgTrpAspGluTyrLeuGluLysGluSerAlaGluTh 400
Db 1140 ACATGAGTTCAAGGAACATTCAGATGGATGATCTTGAGAAAAGAAATCTCCAGAAAC 1199
QY 400 rleupProLeuAspPheLysPheMetProSerGlnGluArgLeuAspLysPheLysVal 420
Db 1200 CCTACCGCTTGACTGTTCAAGCAATGCCCTTCCCAAGAGAGATTAAGCAAAATTAAGGT 1259
QY 420 lIleLeuIleSerLysArgValGlyLeuArgLeuGlnAlaAlaAspMetCysGluAsnG 440
Db 1260 AATTCGATTTCCAAACGGGTAGACTAGCCCTTGAAAGCTGCGACATGTGTAAATCA 1319
QY 440 nPheIleCysProAlaThrValLysSerValHisGlyArgLeuIleAsnValAsnPheAs 460
Db 1320 GTTTATTTGTCAGCTACAGTAAATCACTGATGAGAACATGATGATCAATTTCA 1379
QY 460 pGlyTrpAspGluGlnPheAspGluLeuTyrAspValAspSerHisAspIleLeuProI 480
Db 1380 CGGCTGGAGTGAAGAAATTTGATGAACTGATGATGAGACTCCACAGATATTCTACCGAT 1439
QY 480 eglTrpCysGluAlaHisSerTyrValLeuGlnProProLysLysTyrAsnTyr 498
Db 1440 AGAGTGTGTAAGACGCACAGTTATGTTTACAACTCCGAAAAAATTAACACTAT 1494

RESULT 5
US-09-872-523-76

; Sequence 76, Application US/09872523

; Patent No. US20020137906A1

; GENERAL INFORMATION:

; APPLICANT: Horvitz, H. Robert

; APPLICANT: Davison, Ewa M.

; APPLICANT: Lu, Xiaowei

; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 76
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-76

Alignment Scores:

Pred. No.:	2,84e-287	Length:	2307
Score:	278.00	Matches:	278
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	55.82%	Indels:	0
DB:	10	Gaps:	0

US-09-872-523-5 (1-498) x US-09-872-523-76 (1-2307)

QY 150 AsnCysIleAspGlyGluIleValGlyGlnThrSerLeuSerProLysPheAspGlyGly 169
Db 1138 AATTCGATGATGGCGAAATCGTCGGCGCAAACTTCCTGCTCCAAATTCGATGAAGG 1197
QY 170 LysAlaLeuLeuSerLysHisArgPheLysValGlyGlnArgLeuGluLeuAsnTyr 189
Db 1198 AAGGCTCTCTAAGCAAGCATCTCTTCAAAAGTTGACAAACGCTTGAACTTAATTAAT 1257
QY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGlyArgArgMetAsn 209
Db 1258 TCCAAATTCATCTGAATATACCGCTAGCCGCAATTCAGAAATATGTGACGACGACAAATGAAT 1317
QY 210 ValSerIleThrLysLysAspPheProGluSerLeuProaspAlaAspAspArgGln 229
Db 1318 GTATCTATCAACAAGAAAGACTTCCGAAATCGCTTCCAGATGCAGATGACGACAGCA 1377
QY 230 ValPheSerSerGlySerGlnTyrTrpIleAspGluGlySerPheIlePheProVal 249
Db 1378 GTCTTAAAGCTCTGATCTCAATATGATAGAGAGAGGAAGCTTTCATATTTCTCGTT 1437
QY 250 GlyPheAlaAlaValAsnGlyTyrGlnLeuAsnAlaLysLysGluTyrIleGluHisThr 269
Db 1438 GCATTTGCACAGATCAATGATATCAACTAAATCCGAAAAGCAATATATTGACACACA 1497
QY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyGluAsnProArgTyrAspSerAspAspVal 289
Db 1498 AATTAATTTGCTCAAGCAATTAATAAATGAGAAAATCCAAATGATGACTCAGACGACGTC 1557
QY 290 ThrPheAspGlnLeuAlaLysAspProIleAspPromeIleTrpArgLysValLysVal 309
Db 1558 ACATTTGATCAATTAACAAAAGATCCAAATTTGATGATGATGATGATGATGATGATGAT 1617
QY 310 GlyGlnLysPheGluLeuIleAspProLeuAlaGlnPheAsnAsnLeuHisValAla 329
Db 1618 GGCAGAAAAGTTTACGCTCATGCAACCCCTTGCTCAGCAATTCATTAACCTCCACGTCGCT 1677
QY 330 SerIleLeuLysPheCysLysThrGluGlyTyrLeuIleValGlyMetAspGlyProAsp 349
Db 1678 TCGATTCCTCAAAATTTGCAAAAGTGAAGATATCTTATTTGGGAATGGATGGTCCAGAT 1737
QY 350 AlaLeuGluAspSerPheProIleHisIleAsnAsnThrPheMetPheProValGlyTyr 369
Db 1738 GCACCTTGAGACAGTTTCTTATTCATATCAATTAATTAATTAATTAATTAATTAATTAAT 1797
QY 370 AlaGluLysTyrAsnLeuGluLeuValIleProAspGluPheLysGlyThrPheArgTrp 389
Db 1798 GCGGAAAAGATATATTTGGAACTTGTCCGCCGAGATGAGTTCAAAAGCAATTCAGATGG 1857
QY 390 AspGluTyrLeuGluLysGluSerAlaGluThrLeuProLeuAspLeuPheLysPromeI 409

Db 1858 GATGAATPACTTGGAGAAAGAAATCTGCACAAACCCACCGCTTGACTTTCAGCCAAATG 1917
Oy 410 Prosergingluatgleuaspplyspshelysvalileuileserlysarval 427
Db 1918 CCTTCCCAAGAGAGATTAGACAAATTTAGGTAATTCGATTTCACAAAGGGTT 1971

RESULT 6
US-09-872-523-77
; Sequence 77, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 77
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-77

Alignment Scores:
Pred. No.: 2,84e-287 Length: 2307
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.82% Indels: 0
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-77 (1-2307)

Oy 150 Asncysileaspplgylulilevaliglylnhrserleuserprolyspshelyspplgyl 169
Db 1138 AATTGCATGTGATGGGAAATGTCGGCCAACTTCGCTGCTCCAAATTCGATCAAGG 1197
Oy 170 Lysalaleuleuserlyshisargphelysvaliglylnatrgleuileuansntyr 189
Db 1198 AAGGCTCTCCTAAGCAACATCGTTTCAAGTTGGCAACGCTTGAACATTAATATAT 1257
Oy 190 Seranserthrgluilearvalalaargliegngluilecysglyarargmetasn 209
Db 1258 TCCAAATTCCTGAAATCGCGGTAGCGGCAATTCAGAAATATGTGGACGACGATGAT 1317
Oy 210 Valserliethrlyslasppheprogluserleuproaspaliaaspaspargln 229
Db 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTTCAGATGCGATGACGACGACAA 1377
Oy 230 Valphesergerlyserglnlytrpilleaspplgylserpshelilepheproval 249
Db 1378 GTCTTTAGCTCTGATCTCAATATTTGATAGACGAGGAGAGCTTCTCATATTTCTCTT 1437
Oy 250 Glyphealaleuileasnglytrglnleuansnialyslysglutylrillegluhisrhr 269
Db 1438 GGATTTGACGACGATCAATGATATCACTAAATGCGAAAGAAAGATATATTGACGACACA 1497
Oy 270 Asnysilealaglnalalelysasnlyglusnproargtyraspseraspaspyal 289
Db 1498 AATAAAATTTGCTCAAGCAATATAAAAAATGAGAAATCCAGATATGATCAGACGACACTC 1557
Oy 290 Thrphesapgluleuialeysaspplleaspprometiletrparglysvalysval 309
Db 1558 ACATTTGATCAATTAAGCAAAAGATCCATGATCCCATGATTTGGAGAAAGTTAAGGTT 1617
Oy 310 Glylnyspshelglnleuileaspproleuialgnglnlnpheasnsnleuilehisvalala 329

Db 1618 GGACAAAAGTTGAGCTCATCGACCCCTTGCTCAGCAATTCATACCTCCAGCTCGCT 1677
Oy 330 Serileuleusphesylsyrthrgluglytrileuilevaliglymetaspplproasp 349
Db 1678 TCGATTTCCAAATTTTGGAAAACGSAAGATATCTTATTGGGAATGATGATGCCAAT 1737
Oy 350 Alaleuagluaspserpheproillehisileasnsnhrphenetpheprovalglytr 369
Db 1738 GCACCTGAAGACAGTTTCTTCAATCATATATACATTAATTAATTCCTCCAGTTGGTAT 1797
Oy 370 Alaglylustranleuileuilevalproproasppllyshelysgllythpheartrp 389
Db 1798 GCGGAAAAGTATTAATTTGAACTTGTCCGCAAGATGAGTTCAAGAACATTCAGATGG 1857
Oy 390 Aspglutryleuagluylsugluseralaglutrhleuproleuaspplleuphelsproket 409
Db 1858 GATGAATPACTTGGAGAAAGAAATCTGCACAAACCCACCGCTTGACTTTCAGCCAAATG 1917
Oy 410 Prosergingluatgleuaspplyspshelysvalileuileserlysarval 427
Db 1918 CCTTCCCAAGAGAGATTAGACAAATTTAGGTAATTCGATTTCACAAAGGGTT 1971

RESULT 7
US-09-872-523-78
; Sequence 78, Application US/09872523
; Patent No. US20020137906A1
; GENERAL INFORMATION:
; APPLICANT: Horvitz, H. Robert
; APPLICANT: Davison, Ewa M.
; APPLICANT: Lu, Xiaowei
; TITLE OF INVENTION: A Tumor Suppressor Pathway in C. Elegans
; FILE REFERENCE: 01997/536002
; CURRENT APPLICATION NUMBER: US/09/872,523
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: US 60/208,802
; PRIOR FILING DATE: 2000-06-02
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 78
; LENGTH: 2307
; TYPE: DNA
; ORGANISM: Caenorhabditis elegans
US-09-872-523-78

Alignment Scores:
Pred. No.: 2,84e-287 Length: 2307
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55.82% Indels: 0
DB: 10 Gaps: 0

US-09-872-523-5 (1-498) x US-09-872-523-78 (1-2307)

Oy 150 Asncysileaspplgylulilevaliglylnhrserleuserprolyspshelyspplgyl 169
Db 1138 AATTGCATGTGATGGGAAATGTCGGCCAACTTCGCTGCTCCAAATTCGATCAAGG 1197
Oy 170 Lysalaleuleuserlyshisargphelysvaliglylnatrgleuileuansntyr 189
Db 1198 AAGGCTCTCCTAAGCAACATCGTTTCAAGTTGGCAACGCTTGAACATTAATATAT 1257
Oy 210 Valserliethrlyslasppheprogluserleuproaspaliaaspaspargln 229
Db 1258 TCCAAATTCCTGAAATCGCGGTAGCGGCAATTCAGAAATATGTGGACGACGATGAT 1317
Oy 230 Valphesergerlyserglnlytrpilleaspplgylserpshelilepheproval 249
Db 1318 GTATCTATCACAAAGAAAGAACTTCCGAAATCGCTTCAGATGCGATGACGACGACAA 1377
Oy 250 Valphesergerlyserglnlytrpilleaspplgylserpshelilepheproval 249
Db 1378 GTCTTTAGCTCTGATCTCAATATTTGATAGACGAGGAGAGCTTCTCATATTTCTCTT 1437

QY	250	GLYPHEALALAVLASNGLYTYRGINLEUASNALALYSLYSGLYTYRILLEGILHISTPR	269
DB	1458	GGATTTCGACGACGTCAATGCGATATCACTAAATGCGAAAAAGCAATATTATTGACACACA	1497
QY	270	ASNLYSILLEALAGNALALELYSASNGLYGUASNPORATYTYRASPERSAPSPVA	289
DB	1498	AATAAAATTTGCTCAAGCAATAAAAAATGAGAAATATCCAAAGATATGACTCAGCGACGTC	1557
QY	290	THRPHASPGINLEUALALALYSASPROTILEASPPROMETILLETYPARGLYVALYLSVA	309
DB	1558	ACATTGTGATCAATTAGCAAAAAGATCCAAATGTATCCCATGATTTGGAGAAAAGTTAAGGTT	1617
QY	310	GLYGINLYSPHEGLILEULILEASPPROLEUALAGINGINLPHASASNPENHISVALA	329
DB	1618	GGACAAAAGTTGACCTCATGACACCCCTTGCGTCAGCAATTAATTAACCTCCACGTCGCT	1677
QY	330	SERILEULYSPHECYSALYSTHRLGULYTYRLEULLEVALGYMETASPGLYPROASP	349
DB	1678	TGGAATTCCAAAATTTTGGCAAAACCTAAGGATATCTTATGTGTGGAAATGATGATGCCAGAT	1737
QY	350	ALALEUGUASPSERPHETPROTILEHISILEASASNPENHISPHENETPHETVALGYTYR	369
DB	1738	GCACCTTGAGACAGGTTTTCCTATTCATATCAATTAATATCAATTAATATGTTCCACGTGGTAT	1797
QY	370	ALAIGULYTYRASNLLEUGULULEVALPROPASPGULPHELYSGLYTHRPHETATGTP	389
DB	1798	GCGGAAAAGTATTAATTTGGAACTTGTCGCCGACATGAGCTTAAAGAACATTCACATG	1857
QY	390	ASPGULYTYRLEUGULYSGUASERIALGUTHRLEUPROLEUASPLEUAPHELYSPROMET	409
DB	1858	GATGATATCTTGGAGAAAAGATCTCAGAAACCTTACCGCTTGACATTGTTCAAGCCAAATG	1917
QY	410	PROSERGINGUARGLEUASPLYSPELYSVALILEULILESERLYSARVAL	427
DB	1918	CCTTCCCAAGAGAGATTAGACAAATTTAAGTAATTTGATTTCCAAACGGGTT	1971
RESULT 8			
US-09-878-574-13281			
: Sequence 13281, Application US/09878574			
: Patent No. US20020110548A1			
: GENERAL INFORMATION:			
: APPLICANT: Byrum, Joseph R.			
: APPLICANT: La Rosa, Thomas J.			
: APPLICANT: Thompson, Michael D.			
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with			
: TITLE OF INVENTION: Plants			
: FILE REFERENCE: 38-21(15401)B			
: CURRENT APPLICATION NUMBER: US/09/878,574			
: CURRENT FILING DATE: 2001-12-21			
: PRIOR APPLICATION NUMBER: 09/333,535			
: PRIOR FILING DATE: 1999-06-14			
: NUMBER OF SEQ ID NOS: 15775			
: SEQ ID NO 13281			
: LENGTH: 234			
: TYPE: DNA			
: ORGANISM: Glycine max			
: OTHER INFORMATION: Clone ID: 701066855H1			
US-09-878-574-13281			
Alignment Scores:			
Pred. No.: 28.1 Length: 234			
Score: 8.00 Matches: 8			
Percent Similarity: 100.00% Conservative: 0			
Best Local Similarity: 100.00% Mismatches: 0			
Query Match: 1.61% Indels: 0			
DB: 10 Gaps: 0			
US-09-872-523-5 (1-498) x US-09-878-574-13281 (1-234)			
QY	412	GINGUATGULEASPLYSPELYS	419
DB	30	CAGGAGAGCTAGACAAAGTTCAAG	53

```

RESULT 9
US-10-027-632-19030/C
Sequence 19030, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Mang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0

US-09-872-523-5 (1-498) x US-10-027-632-19030 (1-437)

QY      217 PheProGluSerLeuProAspAla 224
|||||
Db      262 TTTCCTGAGAGCCTACGAGATGCC 229

RESULT 10
US-10-027-632-226880
Sequence 226880, Application US/10027632
GENERAL INFORMATION:
APPLICANT: Mang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
Polymorphisms in the Human Genome
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0

Alignment Scores:
Pred. No.:      50.9      Length:      437
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.61%      Indels:      0
DB:              13      Gaps:      0

```

```
; SEQ ID NO 226880
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226880

Alignment Scores:
Pred. No.:      67.7      Length:      590
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-226880 (1-590)

OY      159 GlnThrSerLeuSerProLysPhe 166
        |||||||
DB      157 CAACGTCCTTTCACCTAAATTC 180

RESULT 11
US-10-027-632-226881
; Sequence 226881, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 226881
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-226881

Alignment Scores:
Pred. No.:      67.7      Length:      590
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-226881 (1-590)

OY      159 GlnThrSerLeuSerProLysPhe 166
        |||||||
DB      157 CAACGTCCTTTCACCTAAATTC 180

RESULT 12
US-10-027-632-222125
; Sequence 222125, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 222125
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222125

Alignment Scores:
Pred. No.:      69.6      Length:      608
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-222125 (1-608)

OY      6 LysIleValArgAlaasnLysLys 13
        |||||||
DB      66 AAAATGTTAGACGCTAATAAAAA 89

RESULT 13
US-10-027-632-6153
; Sequence 6153, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 6153
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-6153

Alignment Scores:
Pred. No.:      69.7      Length:      609
```

```
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 222125
; LENGTH: 608
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-222125

Alignment Scores:
Pred. No.:      69.6      Length:      608
Score:          8.00      Matches:      8
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    1.61%      Indels:      0
DB:             13      Gaps:          0

US-09-872-523-5 (1-498) x US-10-027-632-222125 (1-608)

OY      6 LysIleValArgAlaasnLysLys 13
        |||||||
DB      66 AAAATGTTAGACGCTAATAAAAA 89

RESULT 13
US-10-027-632-6153
; Sequence 6153, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: fastSeq for Windows Version 4.0
; SEQ ID NO 6153
; LENGTH: 609
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-6153

Alignment Scores:
Pred. No.:      69.7      Length:      609
```

```
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-6153 (1-609)
QY 213 ThrlyslvAspphProcluser 220
DB 377 ACCAAGAGGACTCCCGAGTCT 400

RESULT 14
US-10-027-632-224608
; Sequence 224608, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224608
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-224608

Alignment Scores:
Pred. No.: 70 Length: 611
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-224608 (1-611)
QY 177 ArgPhelysValGlyGlnArgLeu 184
DB 159 AGGTCGAAGTGGGCGAGACTG 182

RESULT 15
US-10-027-632-224609
; Sequence 224609, Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 224609
; LENGTH: 611
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-224609

Alignment Scores:
Pred. No.: 70 Length: 611
Score: 8.00 Matches: 8
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.61% Indels: 0
DB: 13 Gaps: 0

US-09-872-523-5 (1-498) x US-10-027-632-224609 (1-611)
QY 177 ArgPhelysValGlyGlnArgLeu 184
DB 159 AGGTCGAAGTGGGCGAGACTG 182
```

Search completed: September 2, 2003, 23:19:00
Job time : 400 secs

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: September 2, 2003, 19:33:40 ; Search time 5500 Seconds
(without alignments)
3704.181 Million cell updates/sec

Title: US-09-872-523-5

Perfect score: 498
Sequence: 1 MSEFLKIVRANKSDRKLDK.....PIGWCEAHSVYLQPPKKNY 498

Scoring table: OLIGO
Xgapop 60.0 , Xgapext 60.0
Ygapop 60.0 , Ygapext 60.0
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2888711 segs, 20454813386 residues

Word size: 1

Total number of hits satisfying chosen parameters: 5773148

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=x1h
-O/cqnt2.1/USPTO.spool/US0987523/runat.02092003.073001.27220/app-query.fasta.1.647
-DB=GenEmbl -OPMT=fastp -SUFFIX=oli.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=quality -THR_MIN=1 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptio
-NORM=ext -HEAPSIZE=500 -MNLN=0 -MAXLEN=200000000
-USBR=US0987523.@CEN.1.1.3508.@runat.02092003.073001.27220 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=60 -YGAEXT=60 -DELOP=6 -DELEXT=7

Database: GenEmbl:*
1: gb_ba:*
2: gb_bhg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hlg_hum:*
31: em_hlg_inv:*
32: em_hlg_other:*
33: em_hlg_mus:*
34: em_hlg_pin:*
35: em_hlg_rod:*
36: em_hlg_mam:*
37: em_hlg_vrt:*
38: em_sy:*
39: em_hlgo_hum:*
40: em_hlgo_mus:*
41: em_hlgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	278	55.8	30911	3 CER06C7	271266 Caenorhabd
C 2	11	2.2	99298	3 AC090416	AC090416 Caenorhab
C 3	10	2.0	142278	9 AC005177	AC005177 Homo sapi
C 4	10	2.0	182051	9 AC007638	AC007638 Homo sapi
C 5	9	1.8	74224	2 AC021724	AC021724 Homo sapi
C 6	9	1.8	98569	9 AC110778	AC110778 Homo sapi
C 7	9	1.8	110000	2 AC123272_0	AC123272 Rattus no
C 8	9	1.8	115532	2 AC008424	AC008424 Homo sapi
C 9	9	1.8	133322	2 AP005306	AP005306 Oryza sat
C 10	9	1.8	140542	2 AC079917	AC079917 Homo sapi
C 11	9	1.8	142623	2 AP005291	AP005291 Oryza sat
C 12	9	1.8	149425	2 AC013320	AC013320 Homo sapi
C 13	9	1.8	154229	2 AC008856	AC008856 Homo sapi
C 14	9	1.8	157102	2 AC016409	AC016409 Homo sapi
C 15	9	1.8	162649	2 BX323824	BX323824 Danio rer
C 16	9	1.8	166447	2 AC021723	AC021723 Homo sapi
C 17	9	1.8	169624	2 AC024187	AC024187 Homo sapi
C 18	9	1.8	170882	2 AC107839	AC107839 Mus muscu
C 19	9	1.8	171777	2 AC084713	AC084713 Homo sapi
C 20	9	1.8	172442	9 AC117530	AC117530 Homo sapi
C 21	9	1.8	177807	9 AC010411	AC010411 Homo sapi
C 22	9	1.8	181435	2 AC024244	AC024244 Homo sapi
C 23	9	1.8	191754	9 AC021016	AC021016 Homo sapi
C 24	9	1.8	233711	2 AC098070	AC098070 Rattus no
C 25	9	1.8	232053	2 AC097148	AC097148 Rattus no
C 26	8	1.6	330	9 AF103485	AF103485 Homo sapi
C 27	8	1.6	461	11 G75687	G75687 99-62654 Pe
C 28	8	1.6	576	11 HSPH25G5	AL159056 H.sapiens
C 29	8	1.6	657	11 BV075348	BV075348 S21P6643
C 30	8	1.6	795	11 BV036204	BV036204 S21P6012
C 31	8	1.6	826	11 BV067302	BV067302 S21P6686
C 32	8	1.6	872	6 BD146582	BD146582 Primer to
C 33	8	1.6	974	3 AF099109	AF099109 Leishman
C 34	8	1.6	1442	6 BD073001	BD073001 70 human
C 35	8	1.6	1497	8 AB072454	AB072454 Coprinops
C 36	8	1.6	1830	9 HSM800595	AL080093 Homo sapi
C 37	8	1.6	2202	6 BD159393	BD159393 Primer to
C 38	8	1.6	2202	9 AK022163	AK022163 Homo sapi
C 39	8	1.6	2271	9 HUMTLEIT	M99436 Human trans
C 40	8	1.6	2588	9 BC017364	BC017364 Homo sapi
C 41	8	1.6	2719	6 BD155878	BD155878 Primer fo
C 42	8	1.6	2719	9 AK001086	AK001086 Homo sapi
C 43	8	1.6	2738	9 AK095046	AK095046 Homo sapi
C 44	8	1.6	3524	6 AX409557	AX409557 Sequence
C 45	8	1.6	3524	6 BD080538	BD080538 Chemokine

ALIGNMENTS

RESULT 1

CER06C7/c
LOCUS 30911 bp DNA linear INV 21-MAY-2003
DEFINITION Caenorhabditis elegans cosmid R06C7, complete sequence.
ACCESSION 271266
VERSION 271266.1 GI:1279324
KEYWORDS HNG: Adenylysuccinate lyase; Myosin heavy chain; Rat T0AD-64
protein like; Serine/threonine-protein kinase; Zinc finger protein.
SOURCE
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
AUTHORS 1
TITLE none.
JOURNAL Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
MEDLINE Science 282 (5396), 2012-2018 (1998)
PUBMED 99069613
REMARK 9851916
REFERENCE 2 (bases 1 to 30911)
AUTHORS Gardner, A.E.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1996) Nematode Sequencing Project, Sanger
Institute, Hinxton, Cambridge CB10 1SA, England and Department of
Genetics, Washington University, St. Louis, MO 63110, USA. E-mail:
jes@esanger.ac.uk or rwenematode.wustl.edu
REMARK Coding sequences below are predicted from computer analysis, using
JOURNAL predictions from GeneIndex (P. Green, U. Washington), and other
available information.
COMMENT Current sequence finishing criteria for the C. elegans genome
sequencing consortium are that all bases are either sequenced
unambiguously on both strands, or on a single strand with both a
dye primer and dye terminator reaction, from distinct subclones.
Exceptions are indicated by an explicit note.
IMPORTANT: This sequence is not the entire insert of clone R06C7.
It may be shorter because we only sequence overlapping sections
once, or longer because we arrange for a small overlap between
neighbouring submissions.
The true left end of clone R06C7 is at 1 in this sequence. The true
right end of clone R06C7 is at 8718 in
sequence 271261.
The true left end of clone F21C3 is at 30808 in this sequence. The
start of this sequence (1..106) overlaps with the end of sequence
298261.
The end of this sequence (30808..30911) overlaps with the start of
sequence 271261.
For a graphical representation of this sequence and its analysis
see: - [http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?](http://wormbase.sanger.ac.uk/perl/ace/elegans/seq/sequence?name=R06C7)
name=R06C7
IMPORTANT: This sequence is NOT necessarily the entire insert of
the specified clone. It may be shorter because we only sequence
overlapping sections once, or longer because we arrange for a small
overlap between neighbouring submissions.
Location/Qualifiers
1..30911
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="1"
/clone="R06C7"
join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:39258..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))
/gene="C01H6.9"

CDS
join(complement(588..775),complement(395..539),
complement(169..295),complement(298261.1:204..368),
complement(298261.1:105..155),
complement(271258.1:40294..40359),
complement(271258.1:40101..40244),
complement(271258.1:39857..39994),
complement(271258.1:39258..39701),
complement(271258.1:38262..39212),
complement(271258.1:38105..38207),
complement(271258.1:37940..38057),
complement(271258.1:37773..37895))
/gene="C01H6.9"
/standard_name="C01H6.9"
/note="contains similarity to Pfam domain: PF00069
(Eukaryotic protein kinase domain), Score=37.5,
E-value=8.2e-10, N=2
cDNA EST yk42h5.3 comes from this gene"
cDNA EST yk42h5.3 comes from this gene"
/codon_start=1
/product="Hypothetical protein C01H6.9"
/protein_id="CAA95847.1"
/db_xref="GOA:O17985"
/db_xref="GI:3878843"
/db_xref="SPTRMBL:O17985"
/translation="MPKPRKVPISGKKARNFARNRROLSVITTECPIDPEYE
FSKNPDLFESNDNPEFPKSKKPKPPVPPSPNVDVLSDEEDAGOFKYPADA
DEEKIKRIKEAFSLDVSLEVFNNNGIRKINGKHIIHDKPEEAMPAPAKID
PAKMEOROKKPVKOVPEMNSKANOVCFSCSIIADHROKMSKLAPSLF
RVETPORSKTIYAKVAKVOCDEPLAIYSTPIKPVNKTAAOLDTISISPTONA
STPLEKOTAGRRKPSFVFDSLSDSCENTIKKEQIAQKANSADQKETTNIPIRE
VSGRVHSSSLAGSIQQPODLEETALISENLSSTLKTPTIESRNSHONKSLMTQO
VSDVEDVEDVFNKRAETSYELHEDEHRCSDMLDESKYSKKRQKSPETWCQ
GMSVTEMODDDLEALFEIEKENETRKPTTLOPOROPSRHDSINSAMESTLQO
FLEPTGEEFKQYSESRAESRNIPGMTTHNDPSTLPLVLDLSEIDSPMOL
LHVQGEKSKWMDLPSKALDGRVRKLGEGAVESTIWDKQVPAIKIYFEEKDC
NRQYFGEHSEDMOTSDVLYEVYVMEKLSLRBDMANSTPNFIMSAFVWGXFP
KQVLISADWSYDKLESEFTRPDVYSIDQNTLVSANGIALDVPVLESNEFSLI
HQVLISMAAEALAELEHRLHLGNVLDHVELELYTVHGQVVPSTGIRKYNII
FTLSRISGATTVYWDLENDPAIFEGDDPQPEVYREMKRKSNNMKFSRTNMTI
VYIANRLIDTICKPGILLTEKRMEIKVLFPFAEFGSGCESTLNEEFSEYEGPIG
MSTTRO"
complement(join(1315..1480,1528..1715,1747..1783,
1834..2156))
/gene="R06C7.2"
complement(join(1315..1480,1528..1715,1747..1783,
1834..2156))
/gene="R06C7.2"
/standard_name="R06C7.2"
/codon_start=1
/product="Hypothetical protein R06C7.2"
/protein_id="CAA95841.1"
/db_xref="GI:3878836"
/db_xref="SPTRMBL:O21772"
/translation="WFSNVHPDSDPRDCKRYVSQTSOHWYVFLKYSFENLIGKRIRK
NFLHLEAEKRYVLFQSVMLKQDTAKTERRTDITQSLKLANFENNNKRIQDEREN
KRLRLINSLKLVSSPENLPLQILFELMSLELDLHVFHFQATIKSLDPLLF
EPTPMYVANEYAKSGRNLMEALIDAIPEVSGEPAAMKQITISIKYKEGSLIDRR
VNASLIIOSNCSEFNSNE"
join(3063..3227,3274..3854,3899..4631,4680..4855,
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
/gene="R06C7.1"
join(3063..3227,3274..3854,3899..4631,4680..4855,
4899..5074,5117..5463,5510..5749,5790..6097,6143..6254)
/gene="R06C7.1"
/standard_name="R06C7.1"
/note="contains similarity to Pfam domain: PF02170 (ZAP
domain), Score=151.7, E-value=4.2e-42, N=1; PF02171 (P1wi
domain), Score=456.7, E-value=6.3e-134, N=1
cDNA EST yk31a12.5 comes from this gene
cDNA EST yk21g1.3 comes from this gene
cDNA EST yk25b2.3 comes from this gene
cDNA EST yk31a12.3 comes from this gene
cDNA EST yk36g4.3 comes from this gene


```

cDNA EST yk13a4.3 comes from this gene
cDNA EST yk34a4.3 comes from this gene
cDNA EST yk21g1.5 comes from this gene
cDNA EST yk25b2.5 comes from this gene
cDNA EST yk34a6.5 comes from this gene
cDNA EST yk36g4.5 comes from this gene
cDNA EST yk13a4.5 comes from this gene
cDNA EST yk93h11.5 comes from this gene
cDNA EST yk125b7.5 comes from this gene
cDNA EST yk125b7.3 comes from this gene
cDNA EST yk236e4.3 comes from this gene
cDNA EST yk236d11.3 comes from this gene
cDNA EST yk348b1.3 comes from this gene
cDNA EST yk367e6.3 comes from this gene
cDNA EST yk367e6.5 comes from this gene
cDNA EST yk513b8.3 comes from this gene
cDNA EST yk669b4.3 comes from this gene
cDNA EST yk671a8.3 comes from this gene
cDNA EST yk250f12.5 comes from this gene
cDNA EST yk513b8.5 comes from this gene
cDNA EST yk669b4.5 comes from this gene
cDNA EST yk671a8.5 comes from this gene
cDNA EST yk75912.3 comes from this gene
cDNA EST yk75912.5 comes from this gene"
/codon_start=1
/product="Hypothetical protein R06C7.1"
/protein_id="CA95839.1"
/db_xref="GI:3878834"
/translat="NSPMPQHPMPMPVPVAPPGAMTPMPVPADAKLHSTGND
DLCIRKLOQNVNDGAKMYMKPTPEQKMRPDIQNVNGIEVTKETTYHREVAHKA
DLSTKEVETFKKKKEDPYODRDKCCIPLAYAKNEPEFKMGNOIVDGGSTL
YTTVNLFSLELDANGTSKAYFQINGADTDNDLTKLCLSLIETIAPRONSITLSENG
KRADONIEVNNREYQFLELALNQCVRTEKRECFEKGKVFELNATEEGEQRCV
DVGDGKOLPYGLKKTIOFTEGPYRGONNSPLVDGKAFHKEQVIOKLEPITGD
PSGNGLNMTREKAAVYKIGLDCYSTYTNKRHLRIGIESATKTRFELPDGKTSI
AEYADKYKISLOYPANLVYCKDRGNMYPFAELMVTNRNNOVTIPOOTGNSOKT
KCAVLPDYRORMLITNGKNAVITLENELVAGITVYSEPLAVQARELDGKLYVR
SVMSDKRRAPRGPFVPRATYPDILMAATVAGBGRFSIGDYNOLGMEIDCKRKG
WYTKPECEGLSTEKIMTQLEKVAASKYVIMITDDAIVHLKQYKALEQNTMIV
ODMKISKANAVYDKRLMTLENIKNTNAYKLVGLNTAVDSAKSKMDEOLIGVSA
PAGATKYMDNKGHNPQILIGFASNVANHEFVDGVSAGODTMASIEDVYONSID
LEPKNKALPKRIIYRSGASGESHASIIAYEIPARAIHIGYSKEIKLITFVTEH
SYREFPDOLRSGKATEMNPICIVIDMNVTRNACQFPLNGHTTIQGRAPPLTYVL
ADDCAPMDRLEBLTTLCHHNOIVSLSTSIPLIVANEYAKRGDLWLGELTTKPI
EAKESQGERLKEITKEIGYKQYDNLNOKRVNA"
/complement(6739..11408)
/gene="dhp-1"
/complement(join(6739..6889,7899..7963,8094..8340,
8392..8457,8905..9182,9550..9761,9840..10017,10603..10783,
11317..11408))
/gene="dhp-1"

```

Alignment Scores:

```

Pred. No.: 6,07e-309 Length: 30911
Score: 278.00 Matches: 278
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 55,82% Indels: 0
DB: 3 Gaps: 0

```

US-09-872-523-5 (1-498) x CER06C7 (1-30911)

```

OY 150 AsnCysIIeaspgIuileValGlgInTrhSerLeuSerProLysPheaspGluGly 169
Db 17800 AATTGCATTCGATGCGCAAAATCGCGCAAACTTCCTGCTCCAAATTCGATGAAGG 17741
OY 170 LysAlaLeuLeuSerLysHsArgPheLysValGlgInArgLeuGluLeuLeuAsnTrp 189
Db 17740 AAGGCTCTCTAAGCAAGCATGCTTTCAAAGTTGACACAGCTCTTGAACATATTAATTTAT 17681

```

```

OY 190 SerAsnSerThrGluIleArgValAlaArgIleGlnGluIleCysGluArgGmetAsn 209
Db 17680 TCCAAATTCCTGTAATACCGCGTACGCCGCAATTCAGAAATATGTGGACGACGATGAAT 17621
OY 210 ValSerIleThrLysLysAspPheProGluSerLeuProAspAlaAspAspArgGln 229
Db 17620 GATCTATATCACAAGAAAGACTTCCGGAATGCTTCAGATGACATGACGACAGCA 17561
OY 230 ValPheSerSerGlySerGlnTrpIleAspGluGlySerPhePheIlePheProVal 249
Db 17560 GTCTTACGCTGTGATCTCAATATGTGATGACGAGGAGGACTTCTTCATATTTCCGT 17501
OY 250 GlyPheAlaAlaValAsnGlyTrpGlnLeuAsnAlaLysLysGluTrpIleGlnHsThr 269
Db 17500 GGAATTGCGACGACGATCATGATTCAACTAAATCCGAAAAGAAATATATGACACACA 17441
OY 270 AsnLysIleAlaGlnAlaIleLysAsnGlyGluAsnProArgTrpAspSerAspAspVal 289
Db 17440 AATPAAATTCCTCAAGCAATPAAAAATGAGAAAATCCAGATATGACTCAGACGACGTC 17381
OY 290 ThrPheaspGluLeuAlaLysAspProIleAspPrometIleTrpArgLysValLysVal 309
Db 17380 ACATTTGATCAATATAGCAAAAAGATCCCAATGTATCCATGATTTGGAGAAAAGTTAAGTT 17321
OY 310 GlyGluLysPheGluLeuIleAspProLeuAlaGlnIlePheAsnAsnLeuHsValAla 329
Db 17320 GGAACAAAAGTTTATAGCTATGATGACCCCTTGCTCAGCAATTCATTAACCTCCAGCTGCT 17261
OY 330 SerIleLeuLysPheCysLysTrhGluGlyTrpLeuIleValGlyMetAspGlyProAsp 349
Db 17260 TCGATTCCTCAATTTGCAAAACCTGAAGATATCTTATTTGGGAATGATGCTCCAGAT 17201
OY 350 AlaLeuGlnAspSerPheProIleHsIleAsnAsnThrPheMetPheProValGlyTrp 369
Db 17200 GCACCTTGAAACAGCTTTCTTATTCATATCAATTAATACATTTATGTTCCAGTTGGTTAT 17141
OY 370 AlaGluLysTrpAsnLeuGluLeuValProProAspGluPheLysGlyTrhPheArgTrp 389
Db 17140 GCGGAAAAGATATATTTGGAACCTGTTCCGCCAGATGATGATTCAAAGAACATTCACATGG 17081
OY 390 AspGluTrpLeuGluLysGluSerAlaGluThrLeuProLeuAspLeuPheLysPromet 409
Db 17080 GATGAATATCTTGAGAAAGATTCGCAAGAAACCTTACCTTGATCTTCAAGCCATG 17021
OY 410 ProSerGlnGluTrpLeuAspLysPheLysValIleLeuIleSerLysArgVal 427
Db 17020 CTTTCCCAAGAGATTAGACAAATTTAAGTAACTCTGATTTCCAAACGGGTT 16967

```

RESULT 2

```

AC090416/c 99298 bp DNA linear INV 17-FEB-2001
LOCUS Caenorhabditis briggsae cosmid CB042D07, complete sequence.
AC090416
ACCESSION AC090416
VERSION AC090416.1 GI:12958081
KEYWORDS HTG.
SOURCE Caenorhabditis briggsae
ORGANISM Caenorhabditis briggsae
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 99298)
The C. briggsae Genome Sequencing Center.
Washington University Genome Sequencing Project
unpublished
2 (bases 1 to 99298)
REFERENCE
AUTHORS Waterston, R.
TITLE Direct Submision
JOURNAL Submitted (17-FEB-2001) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
COMMENT
Submitted by:
Genome Sequencing Center
Department of Genetics, Washington University,
St. Louis, MO 63110, USA

```

e-mail: jspieth@watson.wustl.edu

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES

Source

Location/Qualifiers

1..99298

/organism="Caenorhabditis briggsae"

/mol_type="genomic DNA"

/strain="Gujarat G16"

/db_xref="taxon:6238"

/clone="CB042007"

/complement(67280..67362)

/product="tRNA-Lys"

/note="codon recognized: AAG"

BASE COUNT

29971 a 19350 c 19512 g 30465 t

ORIGIN

Alignment Scores:

Pred. No.:	1..77	Length:	99298
Score:	11.00	Matches:	11
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	2.21%	Indels:	0
DB:	3	Gaps:	0

US-09-872-523-5 (1-498) x AC090416 (1-99298)

Oy 430 ArgfengluAlaAlaAspMetCysGluAsnGln 440

Db 17516 CGTCTCGAAGCAGCATATGTGGAGAACCA 17484

RESULT 3

AC005177/c

LOCUS Homo sapiens chromosome 17, clone HClT462L7, complete sequence.

AC005177

AC005177.1 GI:3287470

AC005177.1

HTG:

Homo sapiens (human)

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,L., Harris,K., Horton,L., Howland,J.C., Hul,L., Jacotot,L., Kann,L., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nachman,A., Nahf,R., Naylor,J., O'Connor,T., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R., Strange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and Zody,M.

Direct Submission

Submitted (02-JUL-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

4 (bases 1 to 142278)

Bitren,B., Fasman,K., Linton,L., Nusbaum,C., Lander,E., Allen,N., Baker,J., Baldwin,J., Barna,N., Beckery,R., Benn,J., Boatn,C., Boutwell,C., Brown,A., Castle,A., Cerny,J., Cooke,P., Depayre,E., Devon,K., Dewar,K., Donelan,L., Etemadi,S., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Gardyna,S., Gensheimer,S., Geraghty,K., Gilmartin,T., Grant,G., Gardyna,B., Harris,K., Horton,L., Howland,J.C., Hul,L., Jacotot,L., Kann,L., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., Meldrim,J., Molla,M., Morris,W., Morrow,J., Mychaleckyj,J., Nachman,A., Nahf,R., Naylor,J., O'Connor,T., Pavlin,B., Peterson,K., Riley,R., Roberts,D., Rossello,R., Roy,A., Shyam,R., Strange-Thomann,N., Stilwell,J., Stojanovic,N., Stone,C., Strickland,C., Subramanian,A., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A., Wang,B., Wheeler,J., Wu,Y., Ye,W.J., Zhao,J. and Zody,M.

Direct Submission

Submitted (06-JUN-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Jul 2, 1998 this sequence version replaced gi:3287188.

All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html.

Location/Qualifiers

1..142278

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="17"

/map="17"

/clone.lib="CITC Human BAC"

305..476

/rpt_family="MLT1E"

512..704

/rpt_family="MLT1E"

1056..1134

/rpt_family="L2"

complement(1697..5025)

/rpt_family="L1PA2"

complement(5034..5344)

/rpt_family="AluY"

complement(5345..8438)

/rpt_family="L1PA2"

complement(8768..8813)

/rpt_family="Alu-rich"

complement(8923..9038)

/rpt_family="L2"

complement(9145..9193)

/rpt_family="MLT1D"

complement(9194..9504)

/rpt_family="AluSg"

complement(9507..9959)

/rpt_family="MLT1D"

10153..10354

/rpt_family="MER58A"

10672..10976

/rpt_family="L2"

10977..11394

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

```

repeat_region      /rpt_family="WSTa"
11395. .11462
repeat_region      /rpt_family="L2"
complement(11481. .11786)
repeat_region      /rpt_family="Aluuo"
11897. .12011
repeat_region      /rpt_family="L2"
complement(12056. .12230)
repeat_region      /rpt_family="L2"
complement(12527. .12575)
repeat_region      /rpt_family="L2"
12736. .12784
repeat_region      /rpt_family="L1MA9"
12785. .12829
repeat_region      /rpt_family="CA)n"
12832. .12939
repeat_region      /rpt_family="L1MA9"
complement(13173. .13236)
repeat_region      /rpt_family="AT_rich"
complement(13356. .13411)
repeat_region      /rpt_family="(CANA)n"
13432. .13563
repeat_region      /rpt_family="MIR"
16605. .16651
repeat_region      /rpt_family="AT_rich"
complement(17420. .17487)
repeat_region      /rpt_family="L2"
complement(17594. .17863)
repeat_region      /rpt_family="L2"
complement(17845. .17927)
repeat_region      /rpt_family="L2"
complement(17982. .18358)
repeat_region      /rpt_family="L2"
18409. .18442
repeat_region      /rpt_family="AT_rich"
complement(18507. .18531)
repeat_region      /rpt_family="AT_rich"
complement(18534. .18556)
repeat_region      /rpt_family="L1PA5"
complement(19200. .19268)
repeat_region      /rpt_family="(TA)n"
complement(22343. .22363)
repeat_region      /rpt_family="AT_rich"
23186. .23485
repeat_region      /rpt_family="AluSc"
complement(23963. .24001)
repeat_region      /rpt_family="AT_rich"
complement(24260. .24657)
repeat_region      /rpt_family="MLT1B"
24962. .25196
repeat_region      /rpt_family="L1ME"
25382. .25999
repeat_region      /rpt_family="L1ME"
26004. .26194
repeat_region      /rpt_family="L1ME2"
complement(26565. .27016)
repeat_region      /rpt_family="L1ME1"
complement(27250. .27297)
repeat_region      /rpt_family="AT_rich"
complement(27728. .27780)
repeat_region      /rpt_family="L2"
complement(28184. .28481)
repeat_region      /rpt_family="AluSc"
complement(28496. .28793)
repeat_region      /rpt_family="AluSc"
complement(28802. .28901)
repeat_region      /rpt_family="(GA)n"
29017. .29145
repeat_region      /rpt_family="FLAM_A"
29164. .29190
repeat_region      /rpt_family="(GGCA)n"
29394. .29482
repeat_region      /rpt_family="MLT2E"

```

```

repeat_region      29432. .29649
/rpt_family="MLT1D"
repeat_region      29631. .29808
/rpt_family="MLT1E"
repeat_region      complement(29957. .30889)
/rpt_family="L1MA7"
repeat_region      complement(31345. .31587)
/rpt_family="Alu7b"
repeat_region      complement(31759. .32055)
/rpt_family="AluSc"
repeat_region      complement(32171. .32295)
/rpt_family="L2"
repeat_region      complement(32327. .33832)
/rpt_family="L2"
repeat_region      complement(35067. .35113)
/rpt_family="AT_rich"
repeat_region      35499. .35737

Alignment Scores:
Pred. No.:      34.6      Length:      142278
Score:          10.00      Matches:      10
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:          2.01%      Indels:      0
DB:              9      Gaps:      0

US-09-872-523-5 (1-498) x AC005177 (1-142278)

QY      243 Serphei1epheprovalGlypheA 252
Db      130587 TCTTTTTCATCTTCTCCTGTGCGTTCGCC 130558

RESULT 4
AC007638      182051 bp      DNA      linear      PRI 02-DEC-2001
LOCUS      AC007638
DEFINITION      Homo sapiens chromosome 17, clone RP11-515017, complete sequence.
ACCESSION      AC007638
VERSION      AC007638.8      GI:17226706
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 182051)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-515017
Unpublished
2 (bases 1 to 182051)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baker,J., Baldwin,J., Barna,N., Beckery,R., Benn,J., Brown,A.,
Castle,A., Cerny,J., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Depayre,E., Devon,K., Dewar,K.,
Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C.,
Funke,R., Gage,D., Galagan,J., Gardyna,S., Gilbert,D., Grant,G.,
Hagos,B., Heaford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Lehoczyk,D., Lieu,C., Locke,K., Macdonald,P.,
Mardquis,N., McEwan,P., McGurk,A., McKenna,K., McLachlan,J.,
Naylor,J., Niloff,M., O'Connor,T., O'Donnell,P., Pavlin,B.,
Peterson,K., Pollara,V., Riley,R., Roberts,D., Roy,A., Severy,P.,
Stange-Thomann,N., Stojanovic,N., Stone,C., Subramanian,A.,
Tesfaye,S., Torruella-Miller,I., Vassiliev,H., Vo,A., Wagner,A.,
Wheeler,J., Wu,X., Wyman,D., Ye,W.J. and Zody,M.
Direct Submission
Submitted (22-MAY-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 182051)
Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campotiano,A., Chang,J., Chazaro,B.,
Choepeil,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,

```

TITLE JOURNAL

REFERENCE

AUTHORS

```

Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,
Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Marcusi, N., Matthews, C.,
McCarthy, M., McEwan, P., McKernan, K., McSheeters, R., Meldrum, J.,
Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,
Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R.,
Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Triggillo, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission
Submitted (02-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 2, 2001 this sequence version replaced g1:13958524.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L458
Center clone name: 515_O_17

FEATURES
source
-----
Location/Qualifiers
1..182051
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="RP11-515017"
/clone_lib="RPC1-11 Human Male BAC"
complement(1353..1468)
repeat_region
/rpt_family="L2"
complement(1805..2130)
repeat_region
/rpt_family="AluY"
2602..2760
/rpt_family="AluYb"
complement(3353..3656)
repeat_region
/rpt_family="AluX"
complement(4256..4412)
repeat_region
/rpt_family="L2"
4584..6080
/rpt_family="L1MEC"
6084..6439
/rpt_family="AluSg"
6502..6762
/rpt_family="AluX"
6855..7007
/rpt_family="L1MEC"
7008..7302
/rpt_family="AluSg"
7303..7592
/rpt_family="L1MEC"
7593..7697
/rpt_family="MIR"
7894..7983
/rpt_family="MIR"
10083..10103
/rpt_family="MIR3"
11175..11198
/rpt_family="AT-rich"
11941..12977
/rpt_family="AT-rich"
/rpt_family="L2"

repeat_region 13421..13462
/rpt_family="AT-rich"
repeat_region 13657..13961
/rpt_family="AluSg"
repeat_region complement(14025..14134)
/rpt_family="L1MC/D"
complement(14156..14355)
/rpt_family="MIR"
repeat_region 15107..15343
/rpt_family="MIR"
repeat_region complement(15613..15918)
/rpt_family="AluSx"
15856..15885
/note="PCR reads only"
15886..16038
/note="PCR reads only"
complement(16039..16103)
/note="PCR reads only"
complement(16252..16336)
/rpt_family="L2"
complement(16879..17191)
/rpt_family="L2"
complement(17385..17797)
/rpt_family="MSTA"
complement(18004..18306)
/rpt_family="AluSg"
complement(18406..18850)
/rpt_family="L2"
18851..19257
/rpt_family="MER31B"
complement(19258..19308)
/rpt_family="L2"
complement(19543..20261)
/rpt_family="LRR34"
20409..20720
/rpt_family="AluY"
20940..21074
/rpt_family="MIR"
21287..21593
/rpt_family="L2"
complement(21594..21750)
/rpt_family="FRAM"
21751..22415
/rpt_family="L2"
22446..22805
/rpt_family="AluSg/X"
22815..23063
/rpt_family="L2"
complement(23803..23918)
/rpt_family="MIR"
complement(25013..25306)
/rpt_family="AluJo"
complement(25483..25688)
/rpt_family="L2"
25769..26010
/rpt_family="MLT1L"
complement(26304..26420)
/rpt_family="MIR"
26831..26851
/rpt_family="AT-rich"
complement(26852..27023)
/rpt_family="FRAM"
complement(27026..27326)
/rpt_family="AluJo"
complement(27337..27691)
/rpt_family="L1MC3"
27692..28054
/rpt_family="THE1B"
complement(28055..28551)
/rpt_family="L1MC3"
28601..28717
/rpt_family="MIR"
complement(29115..29419)

```



```

* 24887 24986: gap of 100 bp
* 24987 25807: contig of 821 bp in length
* 25808 25907: gap of 100 bp
* 25908 26687: contig of 780 bp in length
* 26688 26787: gap of 100 bp
* 26788 27598: contig of 811 bp in length
* 27599 27698: gap of 100 bp
* 27699 28512: contig of 814 bp in length
* 28513 29425: contig of 813 bp in length
* 29426 29525: gap of 100 bp
* 29526 30301: contig of 776 bp in length
* 30302 30402: gap of 100 bp
* 30402 31185: contig of 784 bp in length
* 31186 32075: contig of 790 bp in length
* 32076 32175: gap of 100 bp
* 32176 32970: contig of 795 bp in length
* 32971 33070: gap of 100 bp
* 33071 33874: contig of 804 bp in length
* 33875 33974: gap of 100 bp
* 33975 34771: contig of 797 bp in length
* 34772 34871: gap of 100 bp
* 34872 35682: contig of 811 bp in length
* 35683 35782: gap of 100 bp
* 35783 36591: contig of 809 bp in length
* 36592 36691: gap of 100 bp
* 36692 37463: contig of 772 bp in length
* 37464 37563: gap of 100 bp
* 37564 38365: contig of 802 bp in length
* 38366 38465: gap of 100 bp
* 38466 39264: contig of 799 bp in length
* 39265 39364: gap of 100 bp
* 39365 40166: contig of 802 bp in length
* 40167 40266: gap of 100 bp
* 40267 41046: contig of 780 bp in length
* 41047 41146: gap of 100 bp
* 41147 41945: contig of 799 bp in length
* 41946 42045: gap of 100 bp
* 42046 42841: contig of 796 bp in length
* 42842 42941: gap of 100 bp
* 42942 43751: contig of 810 bp in length
* 43752 43851: gap of 100 bp
* 43852 44656: contig of 805 bp in length
* 44657 44756: gap of 100 bp
* 44757 45564: contig of 808 bp in length
* 45565 45664: gap of 100 bp
* 45665 46480: contig of 816 bp in length
* 46481 46580: gap of 100 bp
* 46581 47397: contig of 817 bp in length
* 47398 47497: gap of 100 bp
* 47498 48300: contig of 803 bp in length
* 48301 48400: gap of 100 bp
* 48401 49208: contig of 808 bp in length
* 49209 49308: gap of 100 bp
* 49309 50099: contig of 791 bp in length
* 50100 50199: gap of 100 bp
* 50200 50961: contig of 766 bp in length
* 50962 51061: gap of 100 bp
* 51062 51869: contig of 808 bp in length
* 51870 51969: gap of 100 bp
* 51970 52775: contig of 806 bp in length
* 52776 52875: gap of 100 bp
* 52876 53675: contig of 800 bp in length
* 53676 53775: gap of 100 bp
* 53776 54577: contig of 802 bp in length
* 54578 54677: gap of 100 bp
* 54678 55451: contig of 774 bp in length
* 55452 55551: gap of 100 bp
* 55552 56348: contig of 797 bp in length
* 56349 56448: gap of 100 bp
* 56449 57260: contig of 812 bp in length
* 57261 57360: gap of 100 bp

```

```

* 57361 58165: contig of 805 bp in length
* 58166 58265: gap of 100 bp
* 58266 59057: contig of 792 bp in length
* 59058 59157: gap of 100 bp
* 59158 59920: contig of 763 bp in length
* 59921 60021: gap of 100 bp
* 60021 60802: contig of 782 bp in length
* 60803 60902: gap of 100 bp
* 60903 61631: contig of 729 bp in length
* 61632 61731: gap of 100 bp
* 61732 62513: contig of 782 bp in length
* 62513 62613: gap of 100 bp
* 62614 63418: contig of 805 bp in length

```

Alignment Scores:

```

Pred. No.: 277 Length: 74224
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

```

US-09-872-523-5 (1-498) x AC021724 (1-74224)

```

Qy 359 Tleasnsnrrhphmetpmetpval 367
Db 70395 ATAAACACACCTTATGTTCCAGTC 70421

```

RESULT 6

AC110778 98569 bp DNA linear PRI 16-APR-2002

LOCUS Homo sapiens BAC clone RP11-320L7 from 4, complete sequence.

AC110778

VERSION AC110778.4 GI:19909444

KEYWORDS

HTG.

SOURCE

ORGANISM

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WMSC
Web site: <http://genome.wustl.edu/gsc>
Contact: saplings@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0320L07

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION: The RPci-11 human BAC library was made from the blood of one male donor, as described by Osogoeawa, K., Moon, P. Y., Zhao, B., Fritngen, E., Paterno, M., Catanesse, J. J., and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosomes libraries. *Genomics* 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is Rp11-55L3, 2000 bp overlap.
Actual start of this clone is at base position 87259 of Rp11-55L3.
Actual end is at base position 98569 of Rp11-320L7.

Polymorphisms exist between AC110778 and AC106892.

Location/Qualifiers

FEATURES	Location/Qualifiers
source	1. .98569

```

-----Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: saplens@watson.wustl.edu
----- Summary Statistics
Center project name: H_NH0320L07
-----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. John D.
McPherson, Department of Genetics, Washington University, St. Louis
MO. For additional information about the map position of this
sequence, see http://genome.wustl.edu/gsc

SOURCE INFORMATION:
The RPI1-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanese, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
The clone sequenced to the left is RP11-55L3, 2000 bp overlap.
Actual start of this clone is at base position 87259 of RP11-55L3;
actual end is at base position 98569 of RP11-320L7.

Polymorphisms exist between AC110778 and AC106892.

CHROMS
source
Location/Qualifiers
1..98569
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="4"
/map="4"
/clone="RP11-320L7"
/clone_lib="RPCI-11"
118..414
/rpt_family="Alu"
/rpt_family="Alu"
3775..4266
/rpt_family="MALR"
4328..4702
/rpt_family="MALR"
4943..5531
/rpt_family="ERV1"
/rpt_family="ERV1"
6855..6900
/rpt_family="AT_rich"
8091..8266
/rpt_family="MIR"
8525..8892
/rpt_family="MALR"
9400..9442
/rpt_family="AT_rich"
10156..10177
/rpt_family="AT_rich"
10741..10805
/rpt_family="L2"

repeat_region 12271..12478
/rpt_family="Alu"
repeat_region 14082..14132
/rpt_family="GA-rich"
repeat_region 14133..14490
/rpt_family="(GGGA)n"
repeat_region 14491..14684
/rpt_family="GA-rich"
repeat_region 14794..15091
/rpt_family="Alu"
repeat_region 16245..16373
/rpt_family="MIR"
repeat_region 16456..16576
/rpt_family="L2"
repeat_region 18135..18205
/rpt_family="(TTTG)n"
repeat_region 18375..18836
/rpt_family="ERV1"
repeat_region 18957..18991
/rpt_family="AT_rich"
repeat_region 19506..19814
/rpt_family="Alu"
repeat_region 23246..23545
/rpt_family="Alu"
repeat_region 24690..24767
/rpt_family="AT_rich"
repeat_region 24839..25133
/rpt_family="Alu"
repeat_region 25392..25503
/rpt_family="MIR"
repeat_region 25898..26823
/rpt_family="L1"
repeat_region 28099..28199
/rpt_family="(TT)n"
repeat_region 28678..28776
/rpt_family="MIR"
repeat_region 28867..28908
/rpt_family="AT_rich"
repeat_region 28979..29126
/rpt_family="MIR"
repeat_region 29127..29437
/rpt_family="Alu"
repeat_region 29438..29521
/rpt_family="MIR"
repeat_region 30197..30218
/rpt_family="AT_rich"
repeat_region 33309..33331
/rpt_family="AT_rich"
repeat_region 34266..34310
/rpt_family="AT_rich"
repeat_region 34518..34546
/rpt_family="AT_rich"
repeat_region 35972..36227
/rpt_family="MIR"
repeat_region 39103..39193
/rpt_family="AT_rich"
repeat_region 40775..40825
/rpt_family="AT_rich"
repeat_region 41371..41456
/rpt_family="MER2_type"
repeat_region 41518..41614
/rpt_family="AT_rich"
repeat_region 41669..41703
/rpt_family="AT_rich"
repeat_region 41885..41928
/rpt_family="AT_rich"
repeat_region 42029..42926
/rpt_family="L1"
repeat_region 43610..43714
/rpt_family="BC200"
repeat_region 43715..44087
/rpt_family="MALR"
repeat_region 44088..44472

```

```

repeat_region      /rpt_family="MALR"
                    44473..45591
repeat_region      /rpt_family="L1"
                    45786..46233
repeat_region      /rpt_family="MALR"
                    46336..46374
repeat_region      /rpt_family="CRNA-11e-ATA"
                    46564..46621
repeat_region      /rpt_family="AT_rich"
                    46757..46788
repeat_region      /rpt_family="(TC)n"

Alignment Scores:
Pred.: 356
Score: 9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 1.81%
DB: 9
Gaps: 0

US-09-872-523-5 (1-498) x AC110778 (1-98569)
Oy 186 LeuleuAenTYrSerASeThrGlu 194
DB 8364 TTGTTGAATYACTACTACTACTGAA 8338

RESULT 7
AC123272.0
MPCOMMENT
Sequence split into 4 fragments LOCUS AC123272 Accession AC123272
Fragment Name Begin End
AC123272_0 1 110000
AC123272_1 100001 210000
AC123272_2 200001 310000
AC123272_3 300001 369520
LOCUS AC123272 369520 bp DNA linear HTG 08-OCT-2002
DEFINITION Rattus norvegicus clone CH230-308E15, *** SEQUENCING IN PROGRESS
***, 10 unordered pieces.
ACCESSION AC123272
VERSION AC123272.3 GI:23269854
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 369520)
Munry,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Angiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswal,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryan,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,Y., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Devila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Detamo,C., Ding,Y., Dinh,H., Divya,K.,
Diaper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregorjis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,M.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hennandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulik,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpaty,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensheva,L., Louiseged,H., Lozado,R.D., Lu,X., Ma,J.,
Maheshwari,M., Mahindaratne,M., Mahmoud,M., Malloy,K., Mangum,A.,

```

```

Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwachukwu,O., Okwou,G., Olarnpusagoon,A., Pal,S., Pairs,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkuch,C.,
Pioppert,F., Polidexter,A., Popovic,D., Pritus,E., Pu,L.,
Puzo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajz,D.,
Sneed,A., Sodergren,E., Song,X.Z., Sorelle,R., Sosa,J.,
Steinle,M., Strong,R., Sutton,A., Swatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,Y., Villasa,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczek,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
Direct Submission
Unpublished
2 (bases 1 to 369520)
Worley,K.C.
Direct Submission
Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 369520)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (08-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 23, 2002 this sequence version replaced gi:21671645.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence
may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GTVI
Center clone name: CH230-308E15
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 226252 bases at least Q40
Consensus quality: 230304 bases at least Q30
Consensus quality: 232703 bases at least Q20
Estimated insert size: 246893; sum-of-contigs estimation
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbankdraftdata.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.

```


* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 48216: contig of 48216 bp in length
* 48217 48316: gap of unknown length
* 48317 302116: contig of 253800 bp in length
* 302117 302216: gap of unknown length
* 302217 316444: contig of 14228 bp in length
* 316445 316544: gap of unknown length
* 316545 317730: contig of 1186 bp in length
* 317731 317830: gap of unknown length
* 317831 319510: contig of 1680 bp in length
* 319511 321036: gap of unknown length
* 321037 321136: gap of unknown length
* 321137 324733: contig of 3597 bp in length
* 324734 324833: gap of unknown length
* 324834 327945: contig of 3112 bp in length
* 327946 328045: gap of unknown length
* 328046 343427: contig of 15382 bp in length
* 343428 343527: gap of unknown length
* 343528 369520: contig of 25993 bp in length.

FEATURES
source
1.369520
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-308EL5"
misc_feature
48317..49352
/note="wgs-contig"
misc_feature
231258..232287
/note="wgs-contig"
BASE COUNT 76044 a 44227 c 43334 g 70304 t 135611 others
ORIGIN

Alignment Scores:
Pred. No.: 392 Length: 110000
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AC123272_0 (1-110000)

Oy 58 GluGLYValIlePheGluThrValVal 66
|||||
Db 93002 GAGGTCATCTTGTGAGACTGAGTT 93028

RESULT 8
AC008424 115532 bp DNA linear HTG 18-JUL-2000
LOCUS Homo sapiens chromosome 5 clone CTC-259F19, WORKING DRAFT SEQUENCE,
DEFINITION
6 ordered pieces.
AC008424
VERSION AC008424.5 GI:9255971
KEYWORDS HTG: HTGS_PHASE2: HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 115532)
DOE Joint Genome Institute.
Sequencing of Human Chromosome 5
Unpublished
2 (bases 1 to 115532)
DOE Joint Genome Institute.
Direct Submission
Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
On Jul 18, 2000 this sequence version replaced gi:7708847.
-----Genome Center
Center: Joint Genome Institute

Center Code: JGI
Web site: <http://www.jgi.doe.gov>

Project Information
Center Project Name: 300134
Center clone name: CIT-HSPC_299F19

Summary Statistics
Consensus quality: 112143 bases at least Q40
Consensus quality: 114467 bases at least Q30
Consensus quality: 114880 bases at least Q20
Estimated insert size: 119000; pulse field gel estimation
Estimated insert size: 115332; sum-of-contrigs estimation
Quality coverage: 5.73 in Q20 bases; pulse field gel estimation
Quality coverage: 5.92 in Q20 bases; sum-of-contrigs estimation.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contrigs. Gaps between the contrigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

1 3722: contig of 3722 bp in length
* 3723 3822: gap of unknown length
* 3823 70532: contig of 66710 bp in length
* 70533 70633: gap of unknown length
* 70634 76281: contig of 5649 bp in length
* 76282 76381: gap of unknown length
* 76382 94828: contig of 18447 bp in length
* 94829 94928: gap of unknown length
* 94929 107602: contig of 12674 bp in length
* 107603 107702: gap of unknown length
* 107703 115532: contig of 7830 bp in length.

FEATURES
source
1.115532
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTC-299F19"
/clone_lib="Caltech human BAC library C"
BASE COUNT 37053 a 20285 c 20614 g 37079 t 501 others
ORIGIN

Alignment Scores:
Pred. No.: 409 Length: 115532
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AC008424 (1-115532)

Oy 359 IleAsnThrPheMetPheProVal 367
|||||
Db 97280 ATAAACACACCTTTATGTTCCAGTC 97306

RESULT 9
AP005306 133322 bp DNA linear HTG 31-MAY-2002
LOCUS Oryza sativa (japonica cultivar-group) chromosome 2 clone P0030D07,
DEFINITION
*** SEQUENCING IN PROGRESS ***
AP005306
VERSION AP005306.1 GI:21280356
KEYWORDS HTG: HTGS_PHASE2.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
TITLE
1
Sasaki, T., Matsumoto, T. and Katayose, Y.
Oryza sativa nipponbare (GAS) genomic DNA, chromosome 2, PAC
clone: P0030007

JOURNAL
REFERENCE
AUTHORS
TITLE
2
Published Only in Database (2002)
Sasaki, T., Matsumoto, T. and Katayose, Y.
Direct Submission
Submitted (29-MAY-2002) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgrp.dna.affrc.go.jp/
Tel: 81-298-38-7441, Fax: 81-298-38-7468)

COMMENT
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a working draft sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES
source
Location/Qualifiers
1..133322
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="P0030007"

BASE COUNT
38223 a 28234 c 28707 g 38158 t

ORIGIN

Alignment Scores:
Pred. No.: 465 Length: 133322
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0

US-09-872-523-5 (1-498) x AP005306 (1-133322)

OY
264 GUTYRTTGGUHSHTAASLSTLE 272
|||||

Db
26156 GAATATATAGAACACTAATAATA 26130

RESULT 10
AC079917/c 140542 bp DNA linear PRI 09-APR-2002

LOCUS
DEFINITION Homo sapiens chromosome 11, clone RP11-55122, complete sequence.
AC079917
AC079917.6 GI:20087233
HTG.
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS
TITLE
1
(bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C. and Lander, E.
Homo sapiens chromosome 11, clone RP11-55122
Unpublished
2
(bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Beda, F., Boguslavsky, L.,
Boukhalter, B., Brown, A., Burkett, G., Campiano, A., Castle, A.,
Choquet, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Ferreira, P.,
Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M.,
Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L.,
Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Larocque, K.,
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Lieu, C., Liu, G.,

REFERENCE
AUTHORS
TITLE
JOURNAL
1
Macdonald, P., Margulis, N., McCarthy, M., McKernan, P., McKernan, K.,
McPheters, R., Meldrum, J., Meneus, L., Mihova, T., Mieng, V.,
Morrow, J., Murphy, T., Naylor, J., Norman, C.H., O'Connor, T.,
O'Donnell, P., O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Rieback, M., Riley, R.,
Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P.,
Sougnen, C., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (17-SEP-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3
(bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McKernan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Nobou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (25-MAR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
4
(bases 1 to 140542)
Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,
Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L.,
Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J.,
Chazaro, B., Choquet, Y., Colangelo, M., Collins, S., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,
Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,
Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,
Kamat, A., Karatas, A., Kells, C., Larocque, K., Lamazares, R.,
Landers, T., Lehoczy, J., Levine, R., Lindblad-Toh, K., Liu, G.,
Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C.,
McCarthy, M., McKernan, P., McKernan, K., Meldrum, J., Meneus, L.,
Mihova, T., Mieng, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R.,
Nobou, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,
Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,
Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,
Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R.,
Seaman, S., Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,
Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J.,
Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.
Direct Submission
Submitted (09-APR-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 9, 2002 this sequence version replaced g1:19703152.
ALL repeats were identified using RepeatMasker:
smt://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research
Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: LI0944
 Center clone name: 55_1_22

 FEATURES
 Source

	Location/Qualifiers
repeat_region	1..140542
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11"
	/clone="RP11-55122"
	/clone_lib="RPC1-11 Human Male BAC"
repeat_region	162..260
repeat_region	/rpt_family="L1MA4"
repeat_region	261..560
repeat_region	/rpt_family="AluY"
repeat_region	318..323
repeat_region	/note="<30 qual SINGL region"
repeat_region	561..770
repeat_region	/rpt_family="L1MA4"
repeat_region	890..929
repeat_region	/rpt_family="AT_rich"
repeat_region	1055..1157
repeat_region	/rpt_family="CT-rich"
repeat_region	complement(1627..2337)
repeat_region	/rpt_family="L1M4"
repeat_region	1980..1984
repeat_region	/note="<30 qual SINGL region"
repeat_region	2117..2122
repeat_region	/note="<30 qual SINGL region"
repeat_region	2153..2157
repeat_region	/note="<30 qual SINGL region"
repeat_region	2175..2180
repeat_region	/note="single clone coverage"
repeat_region	complement(2338..2630)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(2631..4088)
repeat_region	/rpt_family="L1M4"
repeat_region	complement(4098..4475)
repeat_region	/rpt_family="L1M4"
repeat_region	4518..4627
repeat_region	/rpt_family="(TATG)n"
repeat_region	complement(4631..5487)
repeat_region	/rpt_family="L1ME1"
repeat_region	complement(5719..6081)
repeat_region	/rpt_family="L1MA7"
repeat_region	complement(6082..6390)
repeat_region	/rpt_family="AluJb"
repeat_region	complement(6391..6706)
repeat_region	/rpt_family="L1MA7"
repeat_region	complement(6709..7293)
repeat_region	/rpt_family="MSTB1-int"
repeat_region	complement(7298..7716)
repeat_region	/rpt_family="MSTB1"
repeat_region	complement(7717..8744)
repeat_region	/rpt_family="L1MA7"
repeat_region	complement(8745..9042)
repeat_region	/rpt_family="AluSx"
repeat_region	complement(9043..10284)
repeat_region	/rpt_family="L1MA7"
repeat_region	10336..10380
repeat_region	/rpt_family="L1MD2"
repeat_region	10392..10460
repeat_region	/rpt_family="(TTAA)n"
repeat_region	complement(10461..10858)
repeat_region	/rpt_family="L1ME1"
repeat_region	complement(10885..11227)
repeat_region	/rpt_family="L1ME1"
repeat_region	11254..11318
repeat_region	/rpt_family="(TTTA)n"

repeat_region	complement(11350..12565)
repeat_region	/rpt_family="L1PA3"
repeat_region	complement(12578..12979)
repeat_region	/rpt_family="L1ME1"
repeat_region	complement(13067..13120)
repeat_region	/rpt_family="L1ME1"
repeat_region	complement(13121..13429)
repeat_region	/rpt_family="AluY"
repeat_region	complement(13430..14244)
repeat_region	/rpt_family="L1ME1"
repeat_region	14325..14360
repeat_region	/rpt_family="AT_rich"
repeat_region	14467..14485
repeat_region	/rpt_family="L1MD2"
repeat_region	14486..14792
repeat_region	/rpt_family="AluY"
repeat_region	14793..14919
repeat_region	/rpt_family="L1MD2"
repeat_region	14994..15372
repeat_region	/rpt_family="L1MD2"
repeat_region	complement(15371..15704)
repeat_region	/rpt_family="L1M4"
repeat_region	15984..16425
repeat_region	/rpt_family="L1MD2"
repeat_region	complement(16427..16477)
repeat_region	/rpt_family="MLT2B1"
repeat_region	16478..16503
repeat_region	/rpt_family="(CA)n"
repeat_region	complement(16504..16662)
repeat_region	/rpt_family="MLT2B1"
repeat_region	complement(16683..16883)

Alignment Scores:

Pred. No.:	487	Length:	140542
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	9	Gaps:	0

US-09-872-523-5 (1-498) x AC079917 (1-140542)

Qy 418 PhelysValilleuileSerlyArg 426
 Db 41071 TTTAAGTTATATTATTAAGTAAAGA 41045

RESULT 11

AP005291/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AP005291 142623 bp DNA linear HTG 31-MAY-2002
 Oryza sativa (japonica cultivar-group) chromosome 2 clone
 OJ1282.H11, *** SEQUENCING IN PROGRESS ***.
 AP005291 GI:21280341
 HTG: HTGS_PHASE2.
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
 Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae:
 Ehrhartoideae: Oryzaceae: Oryza.
 1
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Oryza sativa nipponbare(GA3) genomic DNA, chromosome 2, BAC
 clone: OJ1282.H11
 Published only in Database (2002)
 2 (bases 1 to 142623)
 Sasaki, T., Matsumoto, T. and Katayose, Y.
 Direct Submission
 Submitted (29-MAY-2002) Takuji Sasaki, National Institute of
 Agricultural Sciences, Rice Genome Research Program, Kannondai
 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
 E-mail: tsasaki@nias.affrc.go.jp, URL: <http://rgp.dna.affrc.go.jp/>,
 Tel: 81-298-38-7441, Fax: 81-298-38-7468
 The nucleotide sequence of this BAC clone was generated by

combining Monsanto and RGE-Japan sequencing data.

NOTE: It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the pieces between them are based on estimates that have provided by the submitter. This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.

* NOTE: This is a 'working draft' sequence.

* This sequence will be replaced

* by the finished sequence as soon as it is available and
* the accession number will be preserved.

FEATURES

source

1. 142623
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/db_xref="taxon:39947"
/chromosome="2"
/clone="OJ1282.H11"

BASE COUNT 40863 a 30430 c 31641 g 39639 t 50 others
ORIGIN

Alignment Scores:

Pred. No.:	493	Length:	142623
Score:	9.00	Matches:	9
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	1.81%	Indels:	0
DB:	2	Gaps:	0

US-09-872-523-5 (1-498) x AP005291 (1-142623)

OY 264 GUTYTTTGGUHHSTHAsnlyStle 272

DB 112946 GAATATATAGAACATCAATAAATA 112920

RESULT 12

AC013320/c

LOCUS AC013320 149425 bp DNA linear HTG 18-NOV-2000
DEFINITION Homo sapiens chromosome 2 clone RP11-36003 map 2, WORKING DRAFT

SEQUENCE, 19 unordered pieces.

ACCESSION AC013320

VERSION AC013320.7 GI:11225433

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 149425)

Homo sapiens chromosome 2, clone RP11-36003

Unpublished

2 (bases 1 to 149425)

REFERENCE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

AUTHORS

JOURNAL

TITLE

http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L3776

Center clone name: 360_O_3

----- Summary Statistics

Sequencing vector: M13; M77815; 12% of reads

Sequencing vector: Plasmid; n/a; 8% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 140887 bases at least Q40

Consensus quality: 144879 bases at least Q30

Consensus quality: 146417 bases at least Q20

Insert size: 228000; agarose-fp

Insert size: 147625; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 5.5 in Q20 ba.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 19 contigs. The true order of the pieces

* is not known and their order in this sequence record is

* arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

1 8144: contig of 8144 bp in length

8145 8244: gap of 100 bp

8245 9350: contig of 1106 bp in length

9351 9450: gap of 100 bp

9451 11335: contig of 1885 bp in length

11336 11435: gap of 100 bp

11436 14325: contig of 2890 bp in length

14326 14425: gap of 100 bp

14426 17305: contig of 2880 bp in length

17306 17405: gap of 100 bp

17406 20166: contig of 2761 bp in length

20167 20266: gap of 100 bp

20267 23474: contig of 3208 bp in length

23475 23574: gap of 100 bp

23575 26403: contig of 2829 bp in length

26404 26503: gap of 100 bp

26504 29899: contig of 3396 bp in length

29900 29999: gap of 100 bp

30000 34929: contig of 4930 bp in length

34930 35029: gap of 100 bp

35030 41463: contig of 6434 bp in length

41464 41563: gap of 100 bp

41564 47890: contig of 6327 bp in length

47891 47990: gap of 100 bp

47991 56092: contig of 8102 bp in length

56093 56192: gap of 100 bp

56193 63063: contig of 6871 bp in length

63064 70819: contig of 7656 bp in length

70820 70920 70919: gap of 100 bp

70920 82482: contig of 11563 bp in length

82483 82582: gap of 100 bp

82583 98645: contig of 16063 bp in length

98646 98745: gap of 100 bp

98746 117782: contig of 15037 bp in length

117783 117882: gap of 100 bp

117883 149425: contig of 31543 bp in length.

FEATURES

source

1. 149425

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="2"

/map="2"

```
/clone="RP11-36003"
/clone.lib="RPC1-11 Human Male BAC"
1..8144
/misc_feature /note="assembly_fragment"
vector_end:SP6
vector_side:left"
8245..9350
/misc_feature /note="assembly_fragment"
9451..11335
/misc_feature /note="assembly_fragment"
11436..14325
/misc_feature /note="assembly_fragment"
14426..17305
/misc_feature /note="assembly_fragment"
17406..20166
/misc_feature /note="assembly_fragment"
20267..22474
/misc_feature /note="assembly_fragment"
23575..26403
/misc_feature /note="assembly_fragment"
26504..29899
/misc_feature /note="assembly_fragment"
30000..34929
/misc_feature /note="assembly_fragment"
35030..41463
/misc_feature /note="assembly_fragment"
41564..47890
/misc_feature /note="assembly_fragment"
47991..56092
/misc_feature /note="assembly_fragment"
56193..63063
/misc_feature /note="assembly_fragment"
63164..70819
/misc_feature /note="assembly_fragment"
70920..82482
/misc_feature /note="assembly_fragment"
82583..98645
/misc_feature /note="assembly_fragment"
98746..117782
/misc_feature /note="assembly_fragment"
117883..149425
/misc_feature /note="assembly_fragment"
vector_side:right"
clone_end:T7
BASE COUNT 36100 a 37729 c 37628 g 36165 t 1803 others
ORIGIN
Alignment Scores:
Pred. No.: 514 Length: 149425
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 2 Gaps: 0
US-09-872-523-5 (1-498) x AC013320 (1-149425)
QY 351 LcUGlUaSpSeRPhcProIeHIsIle 359
Db 20518 TTGGAAGACAGTTTCTATTCACATC 20492
RESULT 13
AC008856/c 154229 bp DNA linear PRI 19-JUL-2001
LOCUS AC008856 Homo sapiens chromosome 5 clone CTD-2178N20, complete sequence.
ACCESSION AC008856
VERSION AC008856.6 GI:14916142
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 154229)
```

```
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 154229)
AUTHORS DOE Joint Genome Institute.
JOURNAL Direct Submission
REFERENCE Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
AUTHORS Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
TITLE 3 (bases 1 to 154229)
JOURNAL Direct Submission
REFERENCE DOE Joint Genome Institute and Stanford Human Genome Center.
AUTHORS Submitted (19-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
TITLE Drive, Walnut Creek, CA 94598, USA
COMMENT On Jul 19, 2001 this sequence version replaced gi:13676976.
Draft Sequence. Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 100% of Sequence;
Estimated Total Number of Errors is 0.
FEATURES
source 1..154229
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="CTD-2178N20"
BASE COUNT 49945 a 27443 c 26935 g 49906 t
ORIGIN
Alignment Scores:
Pred. No.: 529 Length: 154229
Score: 9.00 Matches: 9
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.81% Indels: 0
DB: 9 Gaps: 0
US-09-872-523-5 (1-498) x AC008856 (1-154229)
QY 359 ILASnASnRhPhcMeTcPhcProval 367
Db 18276 ATAAACACACCTTATGTCCACATC 18250
RESULT 14
AC016409 157102 bp DNA linear HTG 03-MAR-2000
LOCUS AC016409 Homo sapiens clone RP11-114022, WORKING DRAFT SEQUENCE, 14
AC016409
ACCESSION AC016409.4 GI:7144774
VERSION AC016409.4
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 157102)
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE Homo sapiens, clone RP11-114022
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 157102)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
TITLE Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Bouhgalter, B.,
JOURNAL Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,
COOKE, P., DeRellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heatford, A., Horton, L.,
Howland, J.C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lehoczky, J., Lien, C., Locke, K., Macdonald, P., Marquis, N.,
McEwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
```

TITLE
JOURNAL
COMMENT

Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.
Direct Submission
Submitted (26-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 3, 2000 this sequence version replaced gi:5980376.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: http://www-seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L1967

Center clone name: L14.O.22

Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150953 bases at least Q40

Consensus quality: 153775 bases at least Q30

Consensus quality: 154865 bases at least Q20

Insert size: 15200; agarose-efp

Insert size: 155802; sum-of-contigs

Quality coverage: 5.0 in Q20 bases; agarose-efp

Quality coverage: 4.9 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

1      872: contig of 872 bp in length
*      1
*      873      972: gap of 100 bp
*      973      4028: contig of 3056 bp in length
*      4029      4128: gap of 100 bp
*      4129      8982: contig of 4854 bp in length
*      8983      9082: gap of 100 bp
*      9083      14834: contig of 5752 bp in length
*      14835      14934: gap of 100 bp
*      14935      21185: contig of 6251 bp in length
*      21186      21285: gap of 100 bp
*      21286      26083: contig of 4798 bp in length
*      26084      26183: gap of 100 bp
*      26184      33624: contig of 7441 bp in length
*      33625      33724: gap of 100 bp
*      33725      40338: contig of 6514 bp in length
*      40339      55110: contig of 14772 bp in length
*      55111      55210: gap of 100 bp
*      55211      68411: contig of 13201 bp in length
*      68412      68511: gap of 100 bp
*      68512      82398: contig of 13887 bp in length
*      82399      82498: gap of 100 bp
*      82499      99701: contig of 17203 bp in length
*      99702      99801: gap of 100 bp
*      99802      124153: contig of 24352 bp in length
*      124154      124253: gap of 100 bp
*      124254      157102: contig of 32849 bp in length.

```

FEATURES

Source

```

1..157102
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Rp11-114022"
/clone_1fb="RPC1-11 Human Male BAC"
1..872
/misc_feature
/note="assembly_fragment"

```

```

misc_feature      clone_end:77
                  vector_side:left"
                  973..4028
                  /note="assembly_fragment"
misc_feature      4129..8982
                  /note="assembly_fragment"
misc_feature      9083..14834
                  /note="assembly_fragment"
misc_feature      14935..21185
                  /note="assembly_fragment"
misc_feature      21286..26083
                  /note="assembly_fragment"
misc_feature      26184..33624
                  /note="assembly_fragment"
misc_feature      33725..40338
                  /note="assembly_fragment"
misc_feature      40339..55110
                  /note="assembly_fragment"
misc_feature      55211..68411
                  /note="assembly_fragment"
misc_feature      68512..82398
                  /note="assembly_fragment"
misc_feature      82499..99701
                  /note="assembly_fragment"
misc_feature      99802..124153
                  /note="assembly_fragment"
misc_feature      124254..157102
                  /note="assembly_fragment"
BASE COUNT      50151 a 28016 c 28164 g 49471 t 1300 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      537      Length:      157102
Score:          9.00      Matches:      9
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      1.81%      Indels:      0
DB:               2      Gaps:      0

```

```

US-09-872-523-5 (1-498) x AC016409 (1-157102)
QY      359      ILeaAsnArThpPheMetPheProVal 367
Db      137976      ATAAACAACCTTATATGTCACAGTC 138002

```

```

RESULT 15      BX323824      162649 bp      DNA      linear      HTG 20-APR-2003
LOCUS      BX323824/C
DEFINITION      Dantio rerio clone DKFZ-37P18, *** SEQUENCING IN PROGRESS ***, 21
UNORDERED PLECES.
ACCESSION      BX323824
VERSION      BX323824.3 GI:30026999
KEYWORDS      HTG; HTGS; PHASEL.
SOURCE      Dantio rerio (zebrafish)
ORGANISM      Dantio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

```

```

REFERENCE      1 (bases 1 to 162649)
AUTHORS      Sims, S.
TITLE      Direct Submission
JOURNAL

```

```

COMMENT      Submitted (19-APR-2003) Wellcome Trust Sanger Institute, Hinxton,
              Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
              zfish-help@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
              On Apr 20, 2003 this sequence version replaced gi:29825483.
              ----- Genome Center
              Center: Wellcome Trust Sanger Institute
              Center code: SC
              Web site: http://www.sanger.ac.uk
              Contact: zfish-help@sanger.ac.uk
              ----- Project Information

```

```
Center project name: ZK37F18
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 149039 bases at least Q40
Consensus quality: 152743 bases at least Q30
Consensus quality: 156027 bases at least Q20
Insert size: 160649; sum-of-contigs
Insert size: 181199; 3.2% error; agarose-fp
Quality coverage: 4.86x in Q20 bases; sum-of-contigs quality
coverage: 4.48x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 21 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1      6362: contig of 6362 bp in length
*      6363      6462: gap of 100 bp
*      6463      14590: contig of 8128 bp in length
*      14591      14690: gap of 100 bp
*      14691      34438: contig of 19748 bp in length
*      34439      34538: gap of 100 bp
*      34539      46632: contig of 12094 bp in length
*      46633      46732: gap of 100 bp
*      46733      53161: contig of 6429 bp in length
*      53162      53261: gap of 100 bp
*      53262      59633: contig of 6372 bp in length
*      59634      59733: gap of 100 bp
*      59734      74713: contig of 14980 bp in length
*      74714      74813: gap of 100 bp
*      74814      77867: contig of 3054 bp in length
*      77868      77967: gap of 100 bp
*      77968      83463: contig of 5496 bp in length
*      83464      83563: gap of 100 bp
*      83564      89435: contig of 5872 bp in length
*      89436      89535: gap of 100 bp
*      89536      94337: contig of 4802 bp in length
*      94338      94437: gap of 100 bp
*      94438      117110: contig of 22673 bp in length
*      117111      117210: gap of 100 bp
*      117211      125419: contig of 8209 bp in length
*      125420      125519: gap of 100 bp
*      125520      129093: contig of 3574 bp in length
*      129094      129193: gap of 100 bp
*      129194      131989: contig of 2796 bp in length
*      131990      132089: gap of 100 bp
*      132090      137224: contig of 5135 bp in length
*      137225      137324: gap of 100 bp
*      137325      142286: contig of 4962 bp in length
*      142287      142386: gap of 100 bp
*      142387      153725: contig of 11339 bp in length
*      153726      153825: gap of 100 bp
*      153826      155913: contig of 2088 bp in length
*      155914      156013: gap of 100 bp
*      156014      160055: contig of 4042 bp in length
*      160056      160155: gap of 100 bp
*      160156      162649: contig of 2494 bp in length.

FEATURES
source
1. .162649
  Location/Qualifiers
    /organism="Danio rerio"
    /mol_type="genomic DNA"
    /db_xref="taxon:7955"
    /clone_lib="DKEY-37F18"
    /clone_lib="DantioKey"
  1. .6362
    /note="assembly_fragment:01786
    fragment_chain:1"
  6463. .14590
    /note="assembly_fragment:00991
```

```
misc_feature      14691..34438
                  /note="assembly_fragment:01163
                  fragment_chain:1"
misc_feature      34539..46632
                  /note="assembly_fragment:02246
                  fragment_chain:1"
misc_feature      46733..53161
                  /note="assembly_fragment:00743
                  fragment_chain:1"
misc_feature      53262..59633
                  /note="assembly_fragment:02049
                  fragment_chain:1"
misc_feature      59734..74713
                  /note="assembly_fragment:00222
                  fragment_chain:1"
misc_feature      74814..77867
                  /note="assembly_fragment:02117
                  fragment_chain:1"
misc_feature      77968..83463
                  /note="assembly_fragment:00362
                  fragment_chain:1"
misc_feature      83564..89435
                  /note="assembly_fragment:00819
                  fragment_chain:1"
misc_feature      89536..94337
                  /note="assembly_fragment:00503
                  fragment_chain:1"
misc_feature      94438..117110
                  /note="assembly_fragment:00155
                  fragment_chain:2"
misc_feature      117211..125419
                  /note="assembly_fragment:02285
                  fragment_chain:2"
misc_feature      125520..129093
                  /note="assembly_fragment:01542
                  fragment_chain:2"
misc_feature      129194..131989
                  /note="assembly_fragment:02189
                  fragment_chain:2"
misc_feature      132090..137224
                  /note="assembly_fragment:00132
                  fragment_chain:3"
misc_feature      137325..142286
                  /note="assembly_fragment:02023
                  fragment_chain:3"
misc_feature      142387..153725
                  /note="assembly_fragment:01215
                  fragment_chain:4"
misc_feature      153826..155913
                  /note="assembly_fragment:02136
                  fragment_chain:4"
misc_feature      156014..160055
                  /note="assembly_fragment:00247"
                  fragment_chain:4"
misc_feature      160156..162649
                  /note="assembly_fragment:01452"
                  fragment_chain:4"

BASE COUNT      47899 a 29936 c 31911 g 50890 t 2013 others
ORIGIN
Alignment Scores:
Pred. No.:      554
Score:          9.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match:    1.81%
DB:             2
Gaps:           0

US-09-872-523-5 (1-498) x BX323824 (1-162649)
QY      395 LygGluserAlaGluThrLeuProLeu 403
Db      107178 AAAGAGTCGACGAACACTTCCTTGC 107152
```

Search completed: September 2, 2003, 22:38:48
Job time : 6693 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 2, 2003, 13:45:39 ; Search time 5682 Seconds

(without alignments)
10778.196 Million cell updates/sec

Title: US-09-872-523-6

Perfect score: 1497

Sequence: 1 atgtcgaattctgaaat.....cgaaaagtaacacattga 1497

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*
1: gb_da:*
2: gb_hlg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_hgc_hum:*
31: em_hgc_inv:*
32: em_hgc_other:*
33: em_hgc_mus:*
34: em_hgc_pln:*
35: em_hgc_rtd:*
36: em_hgc_mam:*
37: em_hgc_vrt:*
38: em_sy:*
39: em_hgc_hum:*
40: em_hgc_mus:*
41: em_hgc_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	1103	73.7	30911	3 CER06C7/c	271266.1 Caenorhabd
c 2	422	28.2	99298	3 AC090416	AC090416 Caenorhab
c 3	139.2	9.3	43096	3 AC024792	AC024792 Caenorhab
c 4	128.8	8.6	31267	2 AC006785	AC006785 Caenorhab
5	60.6	4.0	2878	1 BC020018	BC020018 Mus muscu
6	59	3.9	164681	2 AC101996	AC101996 Mus muscu
7	59	3.9	215435	10 AL662838	AL662838 Mouse DNA
8	58.8	3.9	2960	6 AF285577	AF285577 Mus muscu
9	58.8	3.9	2969	6 AX247680	AX247680 Sequence
10	58.8	3.9	163349	10 AL672233	AL672233 Mouse DNA
11	58.8	3.9	230185	2 AC021579	AC021579 Mus muscu
12	57.4	3.8	3590	6 AX046388	AX046388 Sequence
13	55.8	3.7	2564	9 AK000062	AK000062 Homo sapi
14	55.8	3.7	128328	9 AC005839	AC005839 Homo sapi
15	55.8	3.7	199301	2 BX088589	BX088589 Danio rer
16	55.4	3.7	251144	2 AC133113	AC133113 Rattus no
17	55.4	3.7	258549	2 AC133113	AC133113 Rattus no
18	55.2	3.7	2425	9 AK092756	AK092756 Homo sapi
19	55.2	3.7	3255	9 AK094137	AK094137 Homo sapi
20	55.2	3.7	4101	9 AB058701	AB058701 Homo sapi
21	55	3.7	3385	10 BC030864	BC030864 Mus muscu
22	55	3.7	3399	10 AY237001	AY237001 Mus muscu
23	55	3.7	3725	10 AY237002	AY237002 Mus muscu
24	54.2	3.6	2099	9 AK093255	AK093255 Homo sapi
25	54.2	3.6	2845	9 AK091075	AK091075 Homo sapi
26	54.2	3.6	4534	9 BC039316	BC039316 Homo sapi
27	53.8	3.6	193660	2 AC102262	AC102262 Mus muscu
28	53.4	3.6	1035	6 AX576074	AX576074 Sequence
29	53.4	3.6	2105	9 AB072769	AB072769 Macaca fa
30	53.4	3.6	3250	9 AF149045	AF149045 Homo sapi
31	53.4	3.6	3277	9 BC021252	BC021252 Homo sapi
32	53.4	3.6	3315	9 BC009752	BC009752 Homo sapi
33	53.4	3.6	3321	9 AF149046	AF149046 Homo sapi
34	53.4	3.6	136649	9 HSD635E8	AL110502 Human DNA
35	52.4	3.5	6480	5 FRNINE2	AF146688 Fugu rubr
36	51.8	3.5	3105	6 BD158238	BD158238 Primer fo
37	51.8	3.5	3105	9 AK023124	AK023124 Homo sapi
38	51.4	3.4	3096	9 BC051913	BC051913 Homo sapi
39	51.4	3.4	4130	9 HSA18004	Y18004 Homo sapien
40	51.4	3.4	47167	9 HS757P12	AL031007 Human DNA
41	50.6	3.4	2432	9 AK097052	AK097052 Homo sapi
42	50.6	3.4	2555	6 AX078286	AX078286 Sequence
43	50.6	3.4	3071	9 HSM801538	AL136564 Homo sapi
44	50.6	3.4	3188	6 AX714233	AX714233 Sequence
45	50.6	3.4	3188	9 AK056443	AK056443 Homo sapi

ALIGNMENTS

RESULT 1
LOCUS CER06C7 30911 bp DNA linear 21-MAY-2003
DEFINITION Caenorhabditis elegans cosmid R06C7, complete sequence.
ACCESSION 271266
VERSION 271266.1 GI:1279324
KEYWORDS HTG; Adenylosuccinate lyase; Myosin heavy chain; Rat TOAD-64
protein like; Serine/threonine-protein kinase; Zinc finger protein.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditoidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE 1
AUTHORS none.
TITLE genome sequence of the nematode C. elegans: a platform for

Pred. No. is the number of results predicted by chance to have a

JOURNAL	MEDLINE	PUBMED	REMARK	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
investigating biology. The C. elegans Sequencing Consortium Science 282 (5396), 2012-2018 (1998)	99069613	<PUBMED 9851916	The C.elegans Sequencing Consortium.	2 (bases 1 to 30911)	Gardner,A.E.	Direct Submission Submitted (19-APR-1996) Nematode Sequencing Project, Sanger Institute, Hinxton, Cambridge CB10 1SA, England and Department of Genetics, Washington University, St. Louis, MO 63110, USA. E-mail: jess@genetics.wustl.edu Coding sequences below are predicted from computer analysis, using predictions from GeneFinder (P. Green, U. Washington), and other available information.		
						Current sequence finishing criteria for the C. elegans genome sequencing consortium are that all bases are either sequenced unambiguously on both strands, or on a single strand with both a dye primer and dye terminator reaction, from distinct subclones. Exceptions are indicated by an explicit note.		
						IMPORTANT: This sequence is not the entire insert of clone R06C7. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.		
						The true left end of clone R06C7 is at 1 in this sequence. The true right end of clone R06C7 is at 8718 in		
						sequence Z71261.		
						The true left end of clone F21C3 is at 30808 in this sequence. The start of this sequence (1..106) overlaps with the end of sequence Z98261.		
						The end of this sequence (30808..30911) overlaps with the start of Sequence Z71261.		
						For a graphical representation of this sequence and its analysis see:- http://wormbase.sanger.ac.uk/PertI/ace/elegans/seq/sequence?name=R06C7		
						IMPORTANT: This sequence is NOT necessarily the entire insert of the specified clone. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.		
						Location/Qualifiers		
						1..30911		
						/organism="Caenorhabditis elegans"		
						/mol_type="genomic DNA"		
						/strain="Bristol N2"		
						/db_xref="taxon:6239"		
						/chromosome="1"		
						/clone="R06C7"		
						join(complement(588..775),complement(395..539),		
						complement(169..295),complement(298261.1:204..368),		
						complement(296261.1:105..153))		
						complement(Z71258.1:40294..40359),		
						complement(Z71258.1:40101..40244),		
						complement(Z71258.1:39857..39994),		
						complement(Z71258.1:39256..39701),		
						complement(Z71258.1:38262..39212),		
						complement(Z71258.1:38105..38207),		
						complement(Z71258.1:37940..38057),		
						complement(Z71258.1:37773..37895))		
						/gene="CO1H6.9"		
						join(complement(588..775),complement(395..539),		
						complement(169..295),complement(298261.1:204..368),		
						complement(298261.1:105..153))		
						complement(Z71258.1:40294..40359),		
						complement(Z71258.1:39857..39994),		
						complement(Z71258.1:39256..39701),		
						complement(Z71258.1:38262..39212),		
						complement(Z71258.1:38105..38207),		
						complement(Z71258.1:37940..38057),		
						complement(Z71258.1:37773..37895))		
						/gene="CO1H6.9"		
						/standard_name="CO1H6.9"		
						/note="contains similarity to Pfam domain: PF000069		

(Eukaryotic protein kinase domain) score=37.5, E-value=8.2e-10, N=2

CDNA EST yk42h5.3 comes from this gene
CDNA EST yk42h5.3 comes from this gene
/codon_start=1
/product="Hypothetical protein C01h6.9"
/protein_id="CAA95847.1"
/db_xref="GI:3878842"
/db_xref="GOA:017985"
/db_xref="SPTRMBL:O17985"
/translation="MPPKRIKVPISPGKARNFARNRKRQSSVITTCPIDPEYTE
EKNALAFSEFNDNEPFKSKNNKRPPEPSPVSHVYIISDEDAQOFVPPDDA
DEEETIKTKKEAFSLDDVSLFVNNDGRTVALEENKHI, IHRKQPEARPEAKKID
EKKKKRQKRVKQVPEMNSKRNANQVCOTGRVASCESILADHRONKRSALAPSLF
RVRGCPQKSKMIVPAKHVCCOBRLIAYSTPTRVAKTKKOLDITGSIYVSPKTONA
STPIEKQIAGRPRKSPVVDVSLNDSCEITIKKEEDIAOKSANSOKETTINPIEE
ESGVRHSESLAGSIQDPQDLEETAISNRLSHSTLKTTPLESRNESHQNSKIMTQO
VSDSEVEDVETNRKAETSVTELLHEDHETQDPQDRSSKRSKROKTSPEITQO
GSEJVTMDDODDLEALFEIENEKRIERTTLQPPQRPSSRHDSINAMEMSIQO
PLEDIMGGEERGTQYSESRASERNITPTGMTINHEDSITLPYLEDLSEIDSPKMOI
LHVQGEKERTMDSLPKSLDGRVRVKKLGGVAGVSTPTMGKPAIKIVPEKQGC
NQVREYISEEMQISDVLPVEYIVKELISALDEBANSTPFTIMISAEVYMGKYP
KILLSMWDYKELKESENRPDVYSIDONLILFVANGALIEDVLSNGLSIT
HOLVLSMADEALAEENHRDLHGNVLIDRNGVKKLYTVHGQVGLSTHGLKVNLI
FLTSRSKATTVYMDLENDPAIFEGODDPQDFEYVEMKRNCKSNMKSRRTNLMI
VYIANLI,DTKTCICPKGLLEKRMELKVLFDRAEFGSGCESLTINEEFSDFECPIG
MSTRQ"
complement(join(1315, .1480,1528, .1715,1747, .1783,
1834, .2156))
/gene="R06C7.2"
complement(join(1315, .1480,1528, .1715,1747, .1783,
1834, .2156))
/gene="R06C7.2"
/standard_name="R06C7.2"
/codon_start=1
/product="Hypothetical protein R06C7.2"
/protein_id="CA95841.1"
/db_xref="GI:3878836"
/db_xref="SPTRMBL:O21772"
/translation="MESAHPDPSRDOCKRYRYQTSQSHVYIISDEPNLGRKTIK
NFIHLHAERKELEFQSVMLKDQTAKEKTRQETIQOLSKIANFEENRKKLODBEN
KPKRLNINSELKELVSSPENLPIQILFELSENDELIERFHCATFKFLSOLFLE
FPPRVVANEYAKSGSNLMEALITADIAIEVSGPEARMQIDSIKXKSGFLIDRR
VNSALLIOSNCISFNSNF"
join(3063, .3227,3274, .3854,3899, .4631,4680, .4855,
4899, .5074,5117, .5463,5510, .5749,5790, .6097,6143, .6254)
/gene="R06C7.1"
join(3063, .3227,3274, .3854,3899, .4631,4680, .4855,
4899, .5074,5117, .5463,5510, .5749,5790, .6097,6143, .6254)
/gene="R06C7.1"
/standard_name="R06C7.1"
/note="contains similarity to Pfam domain: PFO2170 (ZAP
domain), score=151.7, E-value=4.2e-42, N=1; PFO2171 (Piwi
domain), score=456.7, E-value=6.3e-134, N=1
CDNA EST yk31a12.5 comes from this gene
CDNA EST yk21g1.3 comes from this gene
CDNA EST yk25b2.3 comes from this gene
CDNA EST yk31a12.3 comes from this gene
CDNA EST yk31a12.3 comes from this gene
CDNA EST yk36g4.3 comes from this gene
CDNA EST yk31a4.3 comes from this gene
CDNA EST yk31a4.3 comes from this gene
CDNA EST yk31a4.3 comes from this gene
CDNA EST yk21g1.5 comes from this gene
CDNA EST yk25b2.5 comes from this gene
CDNA EST yk34a6.5 comes from this gene
CDNA EST yk36g4.5 comes from this gene
CDNA EST yk31a4.5 comes from this gene
CDNA EST yk31a4.5 comes from this gene
CDNA EST yk31a4.5 comes from this gene
CDNA EST yk125b7.5 comes from this gene
CDNA EST yk125b7.5 comes from this gene
CDNA EST yk125b7.3 comes from this gene
CDNA EST yk236e4.3 comes from this gene
CDNA EST yk295d11.3 comes from this gene
CDNA EST yk348b1.3 comes from this gene
CDNA EST yk367e6.3 comes from this gene

```

cDNA EST yk236e4.5 comes from this gene
cDNA EST yk295d11.5 comes from this gene
cDNA EST yk348b1.5 comes from this gene
cDNA EST yk367e6.5 comes from this gene
cDNA EST yk513b8.3 comes from this gene
cDNA EST yk669b4.3 comes from this gene
cDNA EST yk671a8.3 comes from this gene
cDNA EST yk250f12.5 comes from this gene
cDNA EST yk513b8.5 comes from this gene
cDNA EST yk669b4.5 comes from this gene
cDNA EST yk671a8.5 comes from this gene
cDNA EST yk775g12.3 comes from this gene
cDNA EST yk775g12.5 comes from this gene
/codon_start=1
/product="Hypothenelial protein R06C7.1"
/protein_id="CA95839.1"
/db_xref="GI:3878834"
/db_xref="SPTREMBL:Q21770"
/translacion="MSPHPQPHPPMPMPMPVTPAPGAMTPMPVPADAKLHQSTGN
DACIKRLQQLNVEDGAKMYKPTPEKMGKRPVDIQTVFQETKETTVAHRYHAKA
DLSTKEVETTKKGEDEVYDNRDKCNLEPLAVENKPEFMKDNQIVYQSGSTL
YTTVNLSELDANGTSKRVQINGADNDIKLPLSLEIYAPRNSITTSSENG
KRTADONIEVNNREYTOFELALNHCVRSTNRGCEHGVTFVLANTEGFDQDVC
DVGDGQLYPLGLKTTIQIEGPGYGRGNNSLVIDGKAAFHKEQTVIQKLFITGD
PSGLNNMTRREKAAAYKIGDCYSTYNNRRLRIEGLFHESTATKTEFLPDGTCSTI
AEYYADKXKISLQYNNANLVCKDRGNNTYPAELMTYSNORVTLPOQGNOSOKT
KCAVALPDVROHMITGNKAVNITLENELVALGIKVYSPLWVAQAEIDGKELVGR
SVMDGKMRAPPGMPYKPAVPDPAAYVANGRCFSIDVNOLYGMPITDCKKRG
MYIKPRETELSTYETKIMTOLKRYVAAKSCYVIMITDDALVHLKQKALEQRTMYLV
QDMKISKANAIVRDKRLLENLIINKTNVLGGILNVTYSPAKSMDDEQLIIGVSYA
PPAGTKYMDMKNKHLNPQIIIGFASNAVNEHFGDPLAAGQDTMASIEDVONSID
LPEKNRKALEPKRIIYRSGASEGSHSILAYEIPLAAILHGSKEIKLFIYVTKH
SYFFRDLRSGSKATEMNIPGIVLDNATNPACQFPLGHGTTLOGCTAKTPLYTL
ADCKAPMDRLSELTFLCHHOIVLSISTSIPTPIYANEXARKRDLMLSLTKGFI
EAKESGGERLKEFTKEIGYKOTDLNQRVNA"
/gene="dhp-1"
/complement(6739..11408)
CDS
complement(join(6739..6889,7899..7963,8094..8340,
8392..8457,8905..9182,9550..9761,9840..10017,10603..10783,
11317..11408))
/gene="dhp-1"
Query Match 73.7%; Score 1103; DB 3; Length 30911;
Best Local Similarity 81.3%; Pred. No. 1.5e-240;
Matches 1497; Conservative 0; Mismatches 0; Indels 344; Gaps 5;
Qy 1 ATGTCTGAATTTTGTGAAAATTTGTCAGAGCTAACAAAAATCGACAGAAAACCTCGATAG 60
Db 18471 ATGTCTGAATTTTGTGAAAATTTGTCAGAGCTAACAAAAATCGACAGAAAACCTCGATAG 18412
Qy 61 ACCTACTTGGGAAATCTCTATTACATCACTTCGAGAAAGAAAAATCTTTCATTCGA 120
Db 18411 ACCTACTTGGGAAATCTCTATTACATCACTTCGAGAAAGAAAAATCTTTCATTCGA 18352
Qy 121 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAACGATGCGTGAAGAAAGATT 180
Db 18351 GTTGAAGCATTCATCGTACCTTACAGTTAATTTTAAACGATGCGTGAAGAAAGATT 18292
Qy 181 ATC-----TTGGAACA 192
Db 18291 ATGTGTGATTCATATTGTTGTAATTCGTTTAAATATCAATTTTGTATTCGAAACA 18232
Qy 193 GTGTCCATGATTATGACAGAACTGCGATTTCATTCAGTCAAGATGTTTGGACGAAAT 252
Db 18231 GTGTCCATGATTATGACAGAACTGCGATTTCATTCAGTCAAGATGTTTGGACGAAAT 18172
Qy 253 GAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGACTGACAGAAATTTGG 312
Db 18171 GAAAAAGTTTGGCGATACAGAGTTCTGGCTCAGTTTATCGAGACTGACAGAAATTTGG 18112
Qy 313 CTCAATATTTTATCGACGATATGTTGGTTGGCAAA----- 350
Db 18111 CTCAATATTTTATCGACGATATGTTGGTTGGCAAAAGTAAAGTTGACGCTGACGCTTT 18052
```

```

Qy 351 ----- 350
Db 18051 TCTACTATTTCTAAATAATATGTTGTTCTGTTACATAAAATTCYAGACAACTATGTATTA 17992
Qy 351 -----CGCGCAANTGATGAT 366
Db 17991 AAATCTGAAACATTTGATATAATGTAATAATTTGAACTTTCACGCGCGCAATGATGAT 17932
Qy 367 CCCAATATGATATAAATTTGATATATGCTCCGCCCTTGCAATCAACGAAGATACCAAAAT 426
Db 17931 CCCAATATGATATAAATTTGATATATGCTCCGCCCTTGCAATCAACGAAGATACCAAAAT 17872
Qy 427 GATATGTAATTTGTAAT----- 447
Db 17871 GATATGTAATTTGTAATGTAATGTTGTTTTTTTCCGAAATTATGTAATATCATCT 17812
Qy 448 -----AATTGATATGATGCGCAAAATCGTCGGCCAACTTCGCTGCTCCAAAT 496
Db 17811 CACAACCTTCAGAAATTTGATGATGCGCAAAATCGTCGGCCAACTTCGCTGCTCCAAAT 17752
Qy 497 TCGATGAAGGAGAGCTCTCTTAAGCAAGCATGCTTTCAAAAGTTGGACACGCTTGAC 556
Db 17751 TCGATGAAGGAGAGCTCTCTTAAGCAAGCATGCTTTCAAAAGTTGGACACGCTTGAC 17692
Qy 557 TATTAATTTATTCGAATTTCTGAAATAGCGGTAGGCGCAATTCAGAAATATGTGAC 616
Db 17691 TATTAATTTATTCGAATTTCTGAAATAGCGGTAGGCGCAATTCAGAAATATGTGAC 17632
Qy 617 GACGAATGAATATATCTATCAAAAAGAAAGACTTCCGCAATCGCTTCAGATCGATG 676
Db 17631 GACGAATGAATATATCTATCTATCAAAAAGAAAGACTTCCGCAATCGCTTCAGATCGATG 17572
Qy 677 ACGACACAAAGTCTTTAGCTCTGATCTCAATATTGATAGACAGAGGAAGCTTCTTCA 736
Db 17571 ACGACACAAAGTCTTTAGCTCTGATCTCAATATTGATAGACAGAGGAAGCTTCTTCA 17512
Qy 737 TATTTCTGTTGATTTTGGACAGCATCAATGATATCAATTAATGCGAAAAAGAAATATA 796
Db 17511 TATTTCTGTTGATTTTGGACAGCATCAATGATATCAATTAATGCGAAAAAGAAATATA 17452
Qy 797 TTGAGCACACAAATAAATTTGCTCAAGCAATAAAAATGAGAAATCCAAATATGACT 856
Db 17451 TTGAGCACACAAATAAATTTGCTCAAGCAATAAAAATGAGAAATCCAAATATGACT 17392
Qy 857 CAGACGACGTCACATTTGATCAATTAACAAAAGATCCAAATGATCCATGATTTGGAGAA 916
Db 17391 CAGACGACGTCACATTTGATCAATTAACAAAAGATCCAAATGATCCATGATTTGGAGAA 17332
Qy 917 AAGTTAAGTTGACAAAGTTTGAAGTCAATGACCCCTGGCTCAGCAATTCATTAAC 976
Db 17331 AAGTTAAGTTGACAAAGTTTGAAGTCAATGACCCCTGGCTCAGCAATTCATTAAC 17272
Qy 977 TCACGTCGCTTCGATTTCTCAAAATTTTGAACAACTGGAAGATATATTGTGGAAATG 1036
Db 17271 TCACGTCGCTTCGATTTCTCAAAATTTTGAACAACTGGAAGATATATTGTGGAAATG 17212
Qy 1037 ATGTCCAGATGCACTTGAAGACAGTTTCTCTATTCATATCAATTAATACATTTATGTTCC 1096
Db 17211 ATGTCCAGATGCACTTGAAGACAGTTTCTCTATTCATATCAATTAATACATTTATGTTCC 17152
Qy 1097 CAGTTGGTATATGCGAAAGATATATTAATTTGAACTTGTCCGCCAGATGAGTTCAAGAA 1156
Db 17151 CAGTTGGTATATGCGAAAGATATATTAATTTGAACTTGTCCGCCAGATGAGTTCAAGAA 17092
Qy 1157 CATTCAGATGGATGAATTAATTTGGAGAAAGATTCGAGAAACCTACCGCTGACTTGT 1216
Db 17091 CATTCAGATGGATGAATTAATTTGGAGAAAGATTCGAGAAACCTACCGCTGACTTGT 17032
Qy 1217 TCAAGCAATATGCTTCCCAAGAGAGATTTAACAATTTAAGTATTTCTGATTTCCAAAC 1276
Db 17031 TCAAGCAATATGCTTCCCAAGAGAGATTTAACAATTTAAGTATTTCTGATTTCCAAAC 16972
```

QY	1277	G	1277	G	1277
Db	16971	GGGTTGTTTAAATCGCTTTGAGATTGTTCACTATTAAATAGTTATTCATTAATGTTTCTT	16912	16912	16912
QY	1278	-----GGTAGGACTACGCGCTTGAAGCTGCTGACATGTTGTGAAAAATCAGTTATTTGTC	1330	1330	1330
Db	16911	GTTTAAAGTAGGACTACGCGCTTGAAGCTGCTGACATGTTGTGAAAAATCAGTTATTTGTC	16852	16852	16852
QY	1331	CAGCTACAGTGAATATGTTTCATATGAAAGACTGATTAATGTTCATTTTCGACGCTGGGATG	1390	1390	1390
Db	16851	CAGCTACAGTGAATATGTTTCATATGAAAGACTGATTAATGTTCATTTTCGACGCTGGGATG	16792	16792	16792
QY	1391	AGAAATTTGATGAACGTATGATGTGGA	1418	1418	1418
Db	16791	AGAAATTTGATGAACGTATGATGTGGA	16733	16733	16733
QY	1419	-----CTCCCATGATATTTCTACCGATAGATGTTGTGGAAGCGC	1456	1456	1456
Db	16731	AATGAAAAATTCATTCATTTTCAGTCCCATATATTTCTACCGATAGATGTTGTGGAAGCGC	16672	16672	16672
QY	1457	ACAGTTATGTTCTACACACCTCCGAAAAAGTACACTATTTGA	1497	1497	1497
Db	16671	ACAGTTATGTTCTACACACCTCCGAAAAAGTACACTATTTGA	16631	16631	16631
RESULT 2	AC090416/c				
LOCUS	AC090416	99298 bp	DNA	linear	INV 17-FEB-2001
DEFINITION	Caenorhabditis briggsae cosmid CB042D07, complete sequence.				
ACCESSION	AC090416				
VERSION	AC090416.1	GI:12958081			
KEYWORDS	HTG.				
SOURCE	Caenorhabditis briggsae				
ORGANISM	Caenorhabditis briggsae				
REFERENCE	1 (bases 1 to 99298)				
AUTHORS	Washington University Genome Sequencing Center.				
TITLE	The C. briggsae Genome Sequencing Project				
REFERENCE	2 (bases 1 to 99298)				
JOURNAL	Waterston, R.				
AUTHORS	Direct Submission				
TITLE	Submitted (17-FEB-2001)				
JOURNAL	University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
COMMENT	Submitted by:				
	Genome Sequencing Center				
	Department of Genetics, Washington University,				
	St. Louis, MO 63110, USA				
	e-mail: jspliehwatson.wustl.edu				

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

FEATURES	Location/Qualifiers
source	1. .99298

| trna | complement(67280..67362) |

BASE COUNT	29971 a	19350 c	19512 g	30465 t
------------	---------	---------	---------	---------

Query Match	28.2%;	Score 422;	DB 3;	Length 99298;
Best Local Similarity	62.5%;	Pred. No. 1.3e-85;		
Matches 689;	Conservative	0;	Mismatches 390;	Indels 23;
				Gaps 1;

OY	356	CAATGAGTATGCCAATATGAGTAAATATGTAATGCTCCGGCTGCAATCAACGAAG	415
Db	18471	CGATGATCCAAATGTATGTGAAGACCAAAATATTAATGCTCCACACCAAAATGTTATGAGCC	18412
OY	416	AATACCAAAATGATATGTAATATATGTAATATTAATTTGATGATGATGGCGAAATCGTGCGCC	475
Db	18411	TGTACCAACATGATTTTGGATAGTTTATGAGAAAAACAATGATGATGAAATGATTTGGAG	18352
OY	476	AAACTTCGCTGTCTCCAAAAATTCGATGAGGAGAGGCTCTCTTAAGCAACATGCTTTTCA	535
Db	18351	AATTTGGGTCTGTCTCCCGTTTACGATGAGAAATAAAGAGAGCTCTCAAGTCTGCTTTTCC	18292
OY	536	AAGTTGGACAACTCTTTGAACATTTAAATTTTCCAATCTCTGTAAGTAATACGCTTAGCGC	595
Db	18291	GCCTTTGACAGCCCTGGAGAGCTTCTCACTACGCACTTCTACAGAAATAGCTGTGCGAC	18232
OY	596	GAATTCAGAAAAATGTGACGACGAGTAATGTATCTATCCAAAGAAAGACTTTCCGG	655
Db	18231	GTGTTCAGGAAGTCTGTGGTCGTCGGATGAGACGTTCTGTGTCACCAAGAAAGACTATTCAG	18172
OY	656	AATCGCTTCCAGATGCAAGATGACGACAGACAAAGTCTTTTAGCTCTGATCTCAATTTTGG	715
Db	18171	GAATCTTCCGAATTCGAGAGAGACGCTCAACTTCAAGGCTCCGAACTCGTTTGTGA	18112
OY	716	TAGACGGAGGAGAGCTTCTCATTTTCTGTTGGATTTGTCACAGTCAATGATATCAAC	775
Db	18111	TCGACGAGGAGAGCTTCTTTGTGTCTCCAGTTTGTTTCTGCTCTTACAAATTAATCTCAAAAT	18052
OY	776	TAAATGCGAAAAAGGAATATATTGTAGACACACAAATAAATTTGCTCAGCAATAAAAATG	835
Db	18051	TGTTGCCACTGAAAGATATATACAGACACCTCCAAATATCTCGCTTGATGAATG	17992
OY	836	GAGAAATTCCAAGTATGACTCAGAGAGACTCACAATTTGATCAATTAGCAAAAGATCCAA	895
Db	17991	GACAGGAGCCAGAAATATCATATATAGACAGACATTTGATTTACCGTTTATCAAGAGAGATG	17932
OY	896	TTGATCCCAATTTTGGAGAAAGTTAAAGTTGGACAAAAGTTTGGATTCATCAGACCCCT	955
Db	17931	TTGTATTAAGAAATTTGGAAAAATATGAAATCGGACTGAAATTTGAGCTCATATGATCCAT	17872
OY	956	TGGCTCAGCAATTCATATACCTCCACGCTCGCTTCGATTCATTCAAATTTTGCAAAACTGAAG	1015
Db	17871	TGGCTCAGAGATTCGAAAACACTCAAGCTGGCATTCGGTTATCTCGTTGTCACACAGCAATG	17812
OY	1016	GATATCTTATTTGTGGAAATGAGATGCTCCAGATGACTTTGAAGACATTTTCCATTTCAAT	1075
Db	17811	GATATATATTTGTGCGAATGATGATGACCCCATATGAAACGATGAGAGATTTTCCACTTCACA	17752
OY	1076	TCATTAATACATTTATGTCCCACTTGGTTATGCGGAAAAGTATTAATTTGGAACCTTGTC	1135
Db	17751	TCGACAACTCTTATGTCTCTCTTGGATATTTGTGACGCAAGATTAATCTGAAACTGAAAC	17692
OY	1136	CGCCAGATGATTCAAAGAAACATTCAGATGGATGATATCTTTGAGAAAAATTCGACG	1195
Db	17691	AOCAGATAGGTTACAGAAATCGTTCAAAATGGGATGAAATCTTGCGCGTGAGAAATGCC	17632
OY	1196	AAACCTTACGCTTGACTTGTTCAGCCCAATGCCCTTCCAGAGAGATTTAGCAATTTTA	1255
Db	17631	AACCACTCAATTTGAACCTTTTATAGCCGACGCCACCTCAGGAGAGACTTGCAACATTTTC	17572
OY	1256	AGGTAAATCTGATTTCCAA-----ACGGTAGAGACTTAGCGCT	1292
Db	17571	AGGTATACATCAATATAATGATGTATAAATCTCTAAATTTTCAAGTCCGAAATGCGTCT	17512
OY	1293	TGAAGCTCTCACAATGTGTGAAAATCAATTTATTTGTCCAGCTACAGTGAATCAAGTTCA	1352
Db	17511	CGAACACAGATATATGTGCGAAACCAATCATATAGTCCGGCAGAAATCTCTCAATTTCA	17455
OY	1353	TGGAAGACTGATTAATGTCAATTTTGCAGCGCTGGGATGAAGAAATTTGATGAATCTATGA	1412
Db	17451	CGAGAGATCAATTAATGTCAATTTTCATTTGATGATGGAGAAACGTTTGCAGAGCTTACGA	17392
OY	1413	TGTGAGTCTCCATGATATTTCTA	1434

Db 17391 TATTAAGTCAGTTTATGCGCTA 17370

RESULT 3

AC024792/c 43096 bp DNA linear INV 22-NOV-2002
LOCUS Caenorhabditis elegans cosmid Y48G1A, complete sequence.
DEFINITION AC024792
AC024792
AC024792.1 GI:7140348
KEYWORDS HTG.
SOURCE Caenorhabditis elegans
ORGANISM Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
1 (bases 1 to 43096)
Waterston, R.

REFERENCE AUTHORS
TITLE Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium
Science 282 (5396), 2012-2018 (1998)
JOURNAL MEDLINE
99069613
PUBMED 9851916

REFERENCE AUTHORS
TITLE Bemis, G., Ryan, E. and Courtney, L.
JOURNAL The sequence of C. elegans cosmid Y48G1A
UNPUBLISHED (2001)
3 (bases 1 to 43096)
REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (01-MAR-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
4 (bases 1 to 43096)
Waterston, R.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (28-MAR-2000) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
5 (bases 1 to 43096)
REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (09-AUG-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
6 (bases 1 to 43096)
Waterston, R.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2001) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
7 (bases 1 to 43096)
Waterston, R.
REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (24-MAY-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
8 (bases 1 to 43096)
Waterston, R.

REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (28-AUG-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA
9 (bases 1 to 43096)
Waterston, R.
REFERENCE AUTHORS
TITLE Direct Submission
JOURNAL Submitted (22-NOV-2002) Department of Genetics, Washington
University, Genome Sequencing Center, 4444 Forest Park Avenue, St.
Louis, MO 63110, USA

COMMENT
Genome Sequencing Center
Department of Genetics, Washington University
St. Louis, MO 63110, USA, and
Sanger Centre, Hinxton Hall
Cambridge CB10 1HQ, England

email: rwnematode.wustl.edu and jess@sanger.ac.uk
NOTICE: This sequence may not be the entire insert of this clone.
It may be shorter because we only sequence overlapping sections
once, or longer because we provide a small overlap between
neighboring submissions.

This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate chemistry
or covered by high quality data (i.e., phred quality >= 30); an
attempt was made to resolve all sequencing problems, such as
compressions and repeats; all regions were covered by sequence from
more than one m13 subclone.

For a graphical representation of this cosmid sequence and its
analysis see:
<http://www.wormbase.org/db/seq/sequence?name=Y48G1A:class=Sequence>

NEIGHBORING COSMID INFORMATION

The 5' cosmid is C53D5, 1000 bp overlap; the 3' cosmid is R119, 300
bp overlap.

NOTES:

Coding sequences below are the result of integration and manual
review of the following data: computer analysis using the program
GeneFinder (P. Green and L. Hillier, personal communication), the
large scale EST projects of Yoji Kohara
(http://www.ddbj.nig.ac.jp/c-elegans/hm/cf_INDEX.html) and The C.
elegans ORFome cloning project (<http://worldb.dfci.harvard.edu/>),
similarity to other proteins from Blastx analyses
(<http://blast.wustl.edu/>), sequence conservation with C. briggsae
using Jim Kent's WABA alignment program (Genome Research
10:1115-1125, 2000), individual C. elegans Genbank submissions,
and personal communications with C. elegans researchers. tRNAs
are predicted using the program tRNAscan-SE (Lowe, T.M. and
Eddy, S.R., 1997, Nucleic Acids. Res., 25, 955-964).
Location/Qualifiers

FEATURES
source

gene
CDS

1..43096
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y48G1A"
complement(4415..8942)
/gene="Y48G1A.5"
complement(join(4415..4718,5328..6007,6665..7978,
8375..8683,8733..8942))
/gene="Y48G1A.5"
/standard_name="Y48G1A.5"
/note="contains similarity to Pfam domains PF03378
(CAS/CSE protein, C-terminus), PF03810 (Importin-beta
N-terminal domain); coded for by the following C. elegans
cDNAs: YK75c11.3, YK75c11.5, YK85b.5, YK19283.3,
YK19283.5, YK225e12.3, YK225e12.5, YK228g12.3, YK228g12.5,
YK381b10.5, YK446b4.3, YK446b4.5, YK484b6.3, YK484b6.5,
YK508b5.5, YK610a2.3, YK610a2.5, YK640b6.3, YK640b6.5,
YK642f4.3, YK642f4.5, YK650d10.3, YK650d10.5, YK732e2.3,
YK732e2.5"
/codon_start=1
/product="Hypothetical protein Y48G1A.5"
/protein_id="AAK84610.1"
/db_xref="GI:15145461"
/db_xref="wormbase:Y48G1A.5"
/translation="MEQICAAQQTLEPDAAIRKGEALRSLQSNPGYIIQLLV
NEOOIAPQIRIAAVALKNEFKRMWGAPEVEMGEDEQFRSMLEAMFYTKSNVQ
ELISNALLIAQRPDEPKWPDVLPYLSKFLNADLNHVASLASMEQIFRKRFPSKS
ALMKELKKCLISTOEPLTLLRNWVGORDOIAGADIAQMLRVLLILAKVYSIC
SQEIPVEYDKDKMMPHRLHYQDAIPQTSNSGSEPTIDELKHEICEIFLYSORY
EEISEFPDITLAVNMLKSTGPOTRTIDYGCALAEPLISVSOYQVYBGHTGSCVL
KTLAENVCVQNLILRQODMELFDEPLDMKIDIGTGVRRRAIDLAIRGACRRFE


```

* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 2189: contig of 2189 bp in length
2202: gap of unknown length
2203 5146: contig of 2944 bp in length
5147 5159: gap of unknown length
5160 9788: contig of 4629 bp in length
9789 9801: gap of unknown length
9802 15045: contig of 5244 bp in length
15046 15058: gap of unknown length
15059 19039: contig of 3981 bp in length
19040 19052: gap of unknown length
19053 25624: contig of 6569 bp in length
25625 33335: gap of unknown length
33336 33348: contig of 7701 bp in length
33349 41470: gap of unknown length
41471 41483: gap of unknown length
41484 59303: contig of 17820 bp in length
59304 59316: gap of unknown length
59317 74066: contig of 14750 bp in length
74067 74078: gap of unknown length
74079 90677: contig of 16598 bp in length
90678 90690: gap of unknown length
90691 113452: contig of 22762 bp in length
113453 113465: gap of unknown length
113466 154662: contig of 41197 bp in length
154663 154675: gap of unknown length
154676 312267: contig of 157592 bp in length.

FEATURES
    source
        1..312267
            /organism="Caenorhabditis elegans"
            /mol_type="genomic DNA"
            /db_xref="taxon:6239"
            /clone="14861"

BASE COUNT  98825 a 56772 c 57406 g 99095 t 169 others

ORIGIN
Query Match      8.6%: Score 128.8; DB 2; Length 312267;
Best Local Similarity 53.3%; Pred. No. 6.5e-19;
Matches 400; Conservative 0; Mismatches 322; Indels 28; Gaps 5;

QY 528 TCCTTCGCAAGTTGGACACAGCTCTTGAATTAATTAATTCACATTTCTAGTAATACG 587
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112122 TCCTTCGCGCTGAATCAACAGCGCTCGACACTACTTGAGAGCCACAGAAATCCG 112181
QY 588 CGTAGCGCAATTCAGAAATATGTGACGACGACAAATGATATCTATCAACAAGAAAGA 647
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112182 TGTCCGACG-ATCCTCAGGATCCTTGCGCCCGCTGATGATGATGTCACCCCTCAGGA 112240
QY 648 CTTTCCGAAATGCTTCAGATGACATGACGACAGACAGACAGTCTTAAGCTCGA---TC 704
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112241 TTATCCGAAAGATCTTCTCTCCCTAGAGGCCAAAGATGCGACAGTTCAGCATAAACGT 112300
QY 705 TCAATATTTGATAGACAGAGGAGCTTCTTCATATTTCTGCTGATTTGCACAGTCAC 764
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112301 CGAGTTTGGGTGGATGAGTCGAGCTTCTTCCGTTTCCGTCGAGATTTCATGATTA 112360
QY 765 TGAATTCACACTAAATGCGAAAAAGAAATATATTTGACACACAATAATTAATTTCTCAAGC 824
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112361 TGGATTAACGACTAAAGCTACGAGAGGTTATTTGGAACATTTCCAGACGATGCC----- 112414
QY 825 AATTAATAAATGAGAAATTCACAGATATGACTCAGACGACGTACATTTGATCAAT--- 881
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112415 -----AGAAAGTTCTGGAGTTATCATATAAAGATGAGTCACCTTTCCAGACGACTTTT 112465
QY 882 ---AGCAAAAGATCCATTCATGATCCCATGATTTTGGAGAAAGTAAAGTTGGACAAAGTT 938
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112466 TCGTGGGAAACCGACATTTTCAGCTGAAAAATTAATTTATTAAGGTGCGCAAAATTT 112525
QY 939 TGAGCTCATGACCCCTTGCGCTCAGCAATTCATTAACCTCCACGTGCTTCGATTCGAA 998
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

```

Db 112526 TGAATTTTGGATTCACACTCAGCAGCACTACCCAGTCATTTTGGCTGCAACAATTCGAAA 112585
QY 999 ATTTTGCAAACGAGAGATATCTATTTGGAGATGATGTCAGATGCATCTTAAGA 1058
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112586 AATATGCAAAACCCCTGGATTTTAAATTTTACCTGATGACAGCAATGCGACATGA 112645
QY 1059 CAGTTTCCCTATTCATATCAATTAATACATTTATGTTCCCGAGTTGGTATCGGAAAAATGA 1118
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112646 ATCTTCCCATTCATATATATATATATTTATGATACCTCGTGGCGCATCTGAGAAATT 112705
QY 1119 TAAATTTGAACTGTTCCGC---CAGATGAGTTCAAGAAGACATTCAGATGAGATGAATA 1175
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112706 CGGAATCAACAGCTGATTCGCTTCAGCAGCAGTACCGCGGAAATTCAGTGGAGAGGTA 112765
QY 1176 CTGAGGAAGAAATTCGCAAAACCCCTACCGCTTGAATTTTCAAGCATTCGCTTCCA 1235
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112766 TTGGAAGAGAACAGCAGCGGAAAAAATTCCTGATGAGATGCTTAAGCGCTTGCCGAGCAA 112825
QY 1236 AGAGACTTATAGCAAAATTTAAGCTAATTCCT 1265
    |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 112826 GGAGAGAGAGCAGCATGTTTGAAGTGGCTTTT 112855

RESULT 5
BC020018 2878 bp mRNA linear ROD 16-APR-2003
LOCUS BC020018
DEFINITION Mus musculus mbt domain containing 1, mRNA (CDNA clone MGC:29000
IMAGE:2646754), complete cds.
ACCESSION BC020018
VERSION BC020018.1 GI:18043655
KEYWORDS MGC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE
    AUTHORS
        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
        Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
        1 (bases 1 to 2878)
        Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
        Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
        Altschul, S.F., Zeeberg, B., Bueltow, K.H., Schaefer, C.F., Bhat, N.K.,
        Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
        Datchenko, L., Marusina, K., Farmer, A.A., Casavant, T.L.,
        Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
        Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
        Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
        Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
        McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
        Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
        Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
        Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
        Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
        Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
        Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
        Butler, J.E., Schein, J.E., Jones, S.J., Skalski, U., Smallus, D.E.,
        Schnerich, A., Schein, J.E., Jones, S.J., Skalski, U., Smallus, D.E.,
        human and mouse cDNA sequences
        Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
        22388257
        12477932
        2 (bases 1 to 2878)
        Strausberg, R.
        Direct Submission
        Submitted (19-DEC-2001) National Institutes of Health, Mammalian
        Gene Collection (MGC), Cancer Genomics Office, National Cancer
        Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
        USA
        NIH-MGC Project URL: http://mgc.ncl.nih.gov
        Contact: MGC help desk
        Email: cgabs-remail.nih.gov
        Tissue Procurement: Gilbert Smith, Ph.D.
        CDNA Library Preparation: Life Technologies, Inc.
        CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
        DNA Sequencing by: Baylor College of Medicine Human Genome

```


* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 164681: contig of 164681 bp in length.

FEATURES
source

Location/Qualifiers
1. .164681

BASE COUNT	51016	a	29608	c	30671	g	53381	t	5	others
ORIGIN	/clone_11b="RPCI-24 Male Mouse BAC"									

Query Match	3.9%;	Score 59;	DB 2;	Length 164681;
Best Local Similarity	55.7%;	Pred. No. 0.0052;		
Matches 113;	Conservative 0;	Mismatches 90;	Indels 0;	Gaps 0;

0y 1276 CCGGTAGACATACGCCCTTGAAGCTGTACATGTGSAATACAGTTATTTGGCAGCT 1335
|||
29695 CCGGTAGACATCAATTGAAAGCTGTAGCTCTCATGGACCACAGCTTATATGTGTAGCC 29757

0y 1336 ACAGTGAATCGATTTCATGGAGACTGTAATAATGTCATTTCAGCGCGGGATGAAGA 1395
Db 2975 ACAGTTCTCGATTATTCACCCCTCTCTTGAGCATACATTTTGATGTGGGAAGAAGC 2981

OY
1396 TTGATGAACCTGATGATGTGGACTCCCATGATTTCACCAGTAGAAGTGTGGAAGC 1455

Db
2891 TATGACCACTGGCTAGACTGTGTGAGTCCCCTGCACCTCTATCTTTAAAGGTGGTGTCAGTTA 2987

Oy	1456	CACAGTATGTCTACAACTCC	1478
		↑	
Db	29875	ACTGGATATCAACTCAACCTCC	29837

RESULT 7
AL662838

DEFINITION	Mouse DNA sequence from clone RP23-378I13 on chromosome 11, complete sequence.
ACCESSION	AF62838
FEATURES	1..378 bp. GC content 47.34%

KEYWORDS	HTG.	SOURCE	ORGANISM
	Mus	musculus (house mouse)	
	Mus	musculus	

REFERENCE
AUTHORS
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 215435)
Mashreghi-Mohammadi, M.
MURID

JOURNAL
Submitted (21-MAY-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humuany@sanger.ac.uk
clonerequest@sanger.ac.uk
No. 32602

- During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation is not annotated as a mutation.

This sequence was finished as follows unless otherwise noted: all corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

Chemistry or covered by high quality data(i.e., phred quality >= 30): an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one aligned subclone. The 147 kb region was covered by at least

assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given

if the feature table with their source databases: Em:, EMBL, Sw:, SWISSPROT, Tr:, TREMBL, Wp:, WORMPEP. Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-378113 is from the RSCG-23 Mouse PAC library constructed by the group of Pleter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBACE3.0.

FEATURES
Source

Location/Qualifiers
1. .215435

```

/clone_lib="RPCI-23"
BASE COUNT      59987 a 43382 c 45793 g 66273 t
ORIGIN

```

Query Match	3.9%	Score 59;	DB 10;	Length 215435;
Best Local Similarity	55.7%	Pred. No. 0.0051;		
Matches 113; Conservative	0;	Mismatches 90;	Indels 0;	Gaps 0;

QY 1276 CCGGTAGACTACGCCCTGACGCTGCATCTGTGAAATCAGTTATTTTCCAGCT 1335
Db 14936 CCGGTAGAAATGAATTAGAACCTGTAAATCTCATGTGAGCCACGGTTATATTTGTACCC 1505

Qy	1395
1336	ACAGTGAATCACTTCATGGAACACTGATAAATCTCAATTTGCAACGGCTGGATGATGAACA
15056	ACAGTACTCGAATTATTCACCCGCTCTTGAGGATACATTGATATGGTTGGAGAAAG
Db	1511

QY 1396 TTTCATGAACTGTATGACTGTGCACCTCCCATGATATTCTAACCGTAGAGTGTGTGAACCG 1455
| | | | |
Db 15116 TATGACCAGTGAGTGTAGACTGTGTGATGCCCTGACCTCATTCTTTGGGTGTGTCAAGTTA 1517

QY	1456	CACAGTATGTTCTACAACTCC	1478
Db	15176	ACTGGATATCACTACAGCTCC	15198

RESULT 8
AF285577

DEFINITION	Mus musculus transcriptional repressor Scml2 (Scml2) mRNA, complete cds.
ACCESSION	AF285577

KEYWORDS	SOURCE	ORGANISM
Mus musculus (house mouse)	Mus	musculus

REFERENCE
AUTHORS
Mammalia: Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2960)
Wang, P.-J., McCarrey, J. R., Yang, F., and Page, D. C.

JOURNAL Submitted (10-JUL-2000) Whitehead Institute, 9 Cambridge Center, Cambridge, MA 02142, USA

FEATURES Location/Qualifiers

```
/organism="Mus musculus"  
/mol_type="mRNA"  
/db_xref="taxon:10090"
```

```
gene      /tissue_type="testis"  
1.2960  
/gene="Scml2"
```

```
/gene="Scm12"  
/note="sex comb on midleg (Drosophila)-like 2"  
/codon_start=1  
/translation="Scm12"
```

```

/protein_id="AAK31956.1"
/db_xref="GI:13603837"

```



```
misc_feature /note="assembly_fragment"  
32227. .43573  
misc_feature /note="assembly_fragment"  
43674. .57577  
misc_feature /note="assembly_fragment"  
57678. .74056  
misc_feature /note="assembly_fragment"  
74157. .116911  
misc_feature /note="assembly_fragment"  
117012. .139343  
misc_feature /note="assembly_fragment"  
139444. .173294  
misc_feature /note="assembly_fragment"  
173395. .214213  
misc_feature /note="assembly_fragment"  
214314. .230185  
misc_feature /note="assembly_fragment"  
clone_end:T7  
vector_slide:right"  
BASE COUNT 70962 a 44539 c 43526 g 69658 t 1500 others  
ORIGIN
```

```
Query Match 3.9%; Score 58.8; DB 2; Length 230185;  
Best Local Similarity 55.3%; Pred. No. 0.0057;  
Matches 114; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
```

```
OY 1279 GTAGACTACGCTTGAAGCTGCTGACATGTGTGAATAATGATTTATTTGCCAGCTACA 1338  
DB 86438 GTTGGAGATGAGATGATGCTGATGAGATAAATCCATTATGATCTGCTGCCACA 86497  
OY 1339 GTGAATGATTCATGAGAGCTGATTAATGTCATTTCCAGCGCTGGAGATGAAGATT 1398  
DB 86498 ATTGAGCTGCTGCTGGAGATCACTTATCACTTTGATGATGAGATGAGACATTT 86557  
OY 1399 GATGAACGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458  
DB 86558 GATATATGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458  
OY 1459 AGTATGCTTACACACTCCGAATAA 1484  
DB 86618 GGAGATGATTTACAGCCACGAGAAA 86643
```

```
RESULT 12  
AX046388 3590 bp DNA linear PAT 24-NOV-2000  
LOCUS AX046388  
DEFINITION Sequence 55 from Patent WO0011168.  
ACCESSION AX046388  
VERSION AX046388.1 GI:11344369  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
```

```
REFERENCE 1  
AUTHORS Lemischka, I. and Moore, K.  
TITLE Genes that regulate hematopoietic blood forming stem cells and uses thereof  
JOURNAL Patent: WO 0011168-A 55 02-MAR-2000;  
Princeton University (US)  
FEATURES  
source location/Qualifiers  
1. 3590  
/organism="Mus musculus"  
/mol_type="genomic DNA"  
/db_xref="taxon:10090"  
BASE COUNT 1101 a 702 c 797 g 990 t  
ORIGIN
```

```
Query Match 3.8%; Score 57.4; DB 6; Length 3590;  
Best Local Similarity 55.2%; Pred. No. 0.016;  
Matches 112; Conservative 0; Mismatches 91; Indels 0; Gaps 0;  
OY 1276 CGGATAGACTACGCTTGAAGCTGCTGACATGTGTGAATAATGATTTATTTGCCAGCT 1335
```

```
DB 1745 CGGTTAGGAATGAATAATGAGAGCTGATGATCTAGAGCCACGGTTAATATGTTAGCC 1804  
OY 1336 ACAGTGAATACGTTCTGAGAGATGATTAATATGATTTGCGAGCTGGATGAAGAA 1395  
DB 1805 ACAGTTACTGCAATTTATTCACACATCTCTGAGATACATTTTATGATGTTGGAGAAAGAG 1864  
OY 1396 TTGATGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1455  
DB 1865 TATGACAGTGGGAGACTGTGAGTCCCTGACCTATCTGTGAGGTGTGTGACAGTTA 1924  
OY 1456 CACAGTTATGTTCTACAACCTCC 1478  
DB 1925 ACTGGATATCAACTACAGCCTCC 1947
```

```
RESULT 13  
AK000062 2564 bp mRNA linear PRI 22-FEB-2000  
LOCUS AK000062  
DEFINITION Homo sapiens cDNA FLJ20053 fis, clone COL00943.  
ACCESSION AK000062  
VERSION AK000062.1 GI:7019904  
KEYWORDS oligo capping; fis (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
```

```
REFERENCE 1 (sites)  
AUTHORS Kawabata, A., Hikiji, T., Kobatake, N., Inagaki, H., Ikema, Y.,  
Okamoto, S., Ohtani, R., Ota, T., Suzuki, Y., Obayashi, M., Nishi, T.,  
Shibahara, T., Tanaka, T., Nakamura, Y., Isogai, T. and Sugano, S.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2564)  
AUTHORS Sugano, S., Suzuki, Y., Ota, T., Obayashi, M., Nishi, T.,  
Shibahara, T., Tanaka, T. and Nakamura, Y.  
TITLE Direct Submission  
JOURNAL Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,  
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,  
Minato-ku, Tokyo 108-8639, Japan (E-mail: cdna1@ims.u-tokyo.ac.jp),  
Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
COMMENT NEDO human cDNA sequencing project supported by Ministry of  
International Trade and Industry of Japan: cDNA full insert  
sequencing: Research Association for Biotechnology: cDNA library  
construction, 5'- & 3'-end one pass sequencing: Department of  
Virology and Human Genome Center, Institute of Medical Science,  
University of Tokyo (partly supported by Science and Technology  
Agency).
```

```
FEATURES  
source location/Qualifiers  
1. 2564  
/organism="Homo sapiens"  
/mol_type="mRNA"  
/db_xref="taxon:9606"  
/clone="COL00943"  
/tissue="colon"  
/clone_lib="COL"  
/note="Cloning vector pHE18SFLJ3"  
169. 1401  
/note="unnamed protein product"  
/codon_start=1  
/protein_id="BA90919.1"  
/db_xref="GI:7019905"  
/translation="MGTCWGDISENVRVNPNTCSLPYKVFYAGIYKLAGYNAALLR  
YEGFENDSGIDFWCNIGSDIHVPGWCAASCKPLVPRPTIOHKTKNKALVLRUTGA  
KTLPPDFSOKVSESMQYPFKPCMRVEVVDKSLCRTLVAVVEVYIGRLRVYSESD  
RTDFFWCMHNSPLLIHNGMSRSIGHRFRSDITKODGFDTPPHLPAKVEVDSGE  
WFKGEMKLEADIDPLNTSTICVATIRKVLADGFLKIGDGSBAAGSMPFCVHATPSI  
PVGFCETNMETLTPPGYTKLPFKMDYLRETSIAAPKLEPKVDVNNRPFVGMUL  
EAVDMEPRILCVATVTRIIHRLRIHFDEWEERYDQWVDCSPDLVPGWCQLTGYQ  
LQPPASCKEIVRKGVL"
```

```
BASE COUNT 832 a 482 c 560 g 690 t  
ORIGIN
```

	Query Match	3.7%:	Score 55.8:	DB 9:	Length 2564:
	Similarity	54.7%:	Pred. No.0.038:		
	Matches 111:	Conservative 0:	Mismatches 92:	Indels 0:	Gaps 0:
QY	1276	CGGGTAGACATACCCCTTGAAGCTGCACATGTGTGAAATCAGTTATTTGTCCAGCT	1335		
Db	1153	CGTGTAGGAATGAATTAGAAGCAGTATGATCTCATGGAGCCACGTTTAAATATGTGTAGCC	1212		
QY	1336	ACAAGTGAATCAGTTTCATGGAAGCTGATTAATGTCAATTTCCAGCGCTGGGATGGAAGAA	1395		
Db	1213	ACAATAACTCGAATTAATTCATCGTCTCTTGGAGTAACTATTTGATGGATGGGAAGAAGAG	1272		
QY	1396	TTTGATTAACGTGTATGATGTGAGACTTCCCATGATATTTACCGATAGAGATGTGGAAGCG	1455		
Db	1273	TATGATCAGTGGGTAGACTGTGATTCACACTGATCTATCTGTAGGGTGGTGTCAATTA	1332		
QY	1456	CACAGTTATGTCTTCAACAACCTCC	1478		
Db	1333	ACTGGATATCAACTACAGACCTCC	1355		

RESULT	14
AC005839/c	
LOCUS	AC005839 128328 bp DNA linear PRI 30-DEC-1996
DEFINITION	Homo sapiens chromosome 17, clone hRFK_481_C_4, complete sequence.
ACCSSION	AC005839
VERSION	AC005839.1 GI:4079626
KEYWORDS	HTc.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1 (bases 1 to 128328)	Bliren,B.,	linton,L.,	Nusbaum,C.	and Lander,E.
	Homo sapiens chromosome 17, clone hRPK.481_C_4			
	Unpublished			
2 (bases 1 to 128328)	Bliren,B.,	linton,L.,	Nusbaum,C.,	Lander,E.,
	Allen,N.,			
	Anderson,M.,			

TITLE Direct Submission
JOURNAL Submitted (21-OCT-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 128328)
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,

TITLE
Direct Submission
Submitted (30-DEC-1998) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT
On Dec 30, 1998 this sequence version replaced gi:40493329.
All repeats were identified using RepeatMasker: Smtt, A.F.A. &
Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
Only the last 128328 base pairs of this clone are being submitted
The remainder overlaps with accession number AC005920 (WICR
project L444).

```
FEATURES
source
location/Qualifiers
1..128328
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"
/clone="hRPK_481_C_4"
/clone_lib="RP11 human BAC library"
repeat_region
304..437
/rpt_family="AluSg"
repeat_region
438..725
/rpt_family="AluY"
repeat_region
728..757
/rpt_family="(CAAA)n"
repeat_region
758..906
/rpt_family="AluSg"
repeat_region
1398..1707
/rpt_family="AluSx"
repeat_region
2026..2323
/rpt_family="AluJo"
repeat_region
2327..2607
/rpt_family="AluY"
repeat_region
2623..2905
/rpt_family="AluJb"
repeat_region
2937..3243
/rpt_family="AluSx"
repeat_region
3363..3454
/rpt_family="FLAM_C"
repeat_region
3455..3468
/rpt_family="(TAGA)n"
repeat_region
3469..3588
/rpt_family="(CTAG)n"
repeat_region
3589..3619
/rpt_family="(TAGA)n"
repeat_region
3620..3904
/rpt_family="AluSp"
repeat_region
3905..3917
/rpt_family="(TAGA)n"
repeat_region
4137..4436
/rpt_family="AluSg"
repeat_region
4446..4576
/rpt_family="AluJo"
repeat_region
5175..5222
/rpt_family="AT_rich"
repeat_region
5223..6021
/rpt_family="AluSx"
repeat_region
6022..7083
/rpt_family="L2"
repeat_region
7084..7395
/rpt_family="AluSg"
repeat_region
7396..7735
/rpt_family="L2"
repeat_region
7736..7941
/rpt_family="AluSg/X"
repeat_region
7945..8245
/rpt_family="AluSp"
repeat_region
8246..8577
/rpt_family="L2"
repeat_region
8578..8887
/rpt_family="AluSx"
repeat_region
8888..9118
/rpt_family="L2"
repeat_region
9119..10019
/rpt_family="AluSx"
```

repeat_region	complement(10027. .10066) /rpt_family="TAAAn"n"
repeat_region	complement(10067. .10196) /rpt_family="FRAM"n"
repeat_region	complement(10563. .10612) /rpt_family="L2" complement(10879. .11191) /rpt_family="AluSg" complement(11324. .11630) /rpt_family="AluSg" complement(12443. .12762) /rpt_family="AluSx" 12842. .13141 /rpt_family="AluSg" 13147. .13434 /rpt_family="AluSg" complement(14308. .14527) /rpt_family="AluDb" 14998. .15291 /rpt_family="AluY" 15300. .15575 /rpt_family="AluSc" 15527. .15536 /note="Single-stranded terminator coverage." 15576. .15596 /rpt_family="(CAAA)n" 15597. .15615 /rpt_family="AluSc" 15625. .15906 /rpt_family="AluSg" 15907. .15934 /rpt_family="(CAAA)n" 15944. .16033 /rpt_family="L1P82" 16036. .16153 /rpt_family="(TA)n" complement(16165. .16459) /rpt_family="AluSx" complement(16536. .16613) /rpt_family="L2" 16723. .16997 /rpt_family="AluSg" 17461. .17779 /rpt_family="AluDb" complement(17976. .17995) /rpt_family="(CAA)n" 18466. .18489 /rpt_family="AT_rich" 18652. .19114 /rpt_family="L1MC5" 19115. .19428 /rpt_family="AluDo" 19429. .19565 /rpt_family="L1MC5" 19576. .19708 /rpt_family="FAM_C" 19777. .19903 /rpt_family="L1ME" 19904. .20203 /rpt_family="AluY" 20204. .20222 /rpt_family="L1ME" 20677. .20802 /rpt_family="HML1" 20996. .21154 /rpt_family="L1MC5" complement(21185. .21272) /rpt_family="L1MC5" 21497. .21805 /rpt_family="AluY" 21811. .21923 /rpt_family="AluSg" 21950. .22149

```

repeat_region      /rpl_family="A1usg"
                   22150..22265
                   /rpl_family="(TA)n"
repeat_region      22429..22554

Query Match      3.7%; Score 55.8; DB 9; Length 128328;
Best Local Similarity 54.7%; Pred. No. 0.029;
Matches 111; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY      1276 CCGGTGAGCACTACGCCCTTGAAAGCTGCTACATGTGTGAAAATCAGTTATTATTGTCACGCT 1335
          ||| ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      119478 CCGGTGAGCAATTAATTAATTAAGCACTACATCTCATGTGAGCCACCTTTAAATGTGTAGCC 119419

QY      1336 ACAGTAAATCAGTTCATGAGGAAGACTGATTAATGTCATTTGACGCGCTGGATGAGAAGAA 1395
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      119418 ACAGTAACTCGAATTAATTCATGCTCTTGTGAGATACATTTTGTGATGGAGGAAGAAAG 119359

QY      1396 TTGTGAGCACTGTATGATGTGGACTGCCCATCTATATTCACGATGAGATGGTGTGAACG 1455
          ||||| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      119358 TATGATCACTGGGTGATGACGTGTAGTCACCTGACCTGTATCTGTAGAGGTGTGTCACCTTA 119299

QY      1456 CACAGTTATGTTCTACACACTCC 1478
          | || | |||| | ||||
Db      119298 ACTGATATCATCACTACAGCTCC 119276

RESULT 15
BX088589      199301 bp      DNA      linear      HTG 15-JAN-2003
LOCUS      BX088589
DEFINITION      Danio rerio clone DKEX-27E7, *** SEQUENCING IN PROGRESS ***, 37
unnumbered pieces.
ACCESSION      BX088589
VERSION      BX088589.1 GI:27764200
KEYWORDS      HTG; HTGS_PHASE1.
SOURCE      Danio rerio (zebrafish)
ORGANISM      Danio rerio
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
              Cypriniformes; Cyprinidae; Danio.
              1 (bases 1 to 199301)
              Burton J.
Direct Submission
Submitted (14-JAN-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfish-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfish-help@sanger.ac.uk
----- Project Information
Center project name: ZK27E7
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 100% of reads
Consensus quality: 182613 bases at least Q40
Consensus quality: 187913 bases at least Q30
Consensus quality: 191719 bases at least Q20
Insert size: 195701; sum-of-contigs
Insert size: 237259; 8.1% error; agarose-fp
Quality coverage: 2.54x in Q20 bases; sum-of-contigs Quality
coverage: 2.29x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 37 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 7924 7923: contig of 7923 bp in length
* 8023: gap of 100 bp

```

8024	10536:	contig of 2513 bp in length	/organism="Danio rerio"
10537	10636:	gap of 100 bp	/mol_type="genomic DNA"
10637	12687:	contig of 2051 bp in length	/db_xref="taxon:7955"
12688	12787:	gap of 100 bp	/clone="DKEY-27E7"
12788	15386:	contig of 2599 bp in length	/clone_lib="DantioKey"
15387	15486:	gap of 100 bp	1. .7923
15487	21116:	contig of 5630 bp in length	/note="assembly_fragment:00057
21117	21216:	gap of 100 bp	fragment_chain:1"
21217	27096:	contig of 5880 bp in length	8024. .10536
27097	27196:	gap of 100 bp	/note="assembly_fragment:00635
27197	31488:	contig of 4292 bp in length	fragment_chain:1"
31489	31588:	gap of 100 bp	10637. .12687
31589	35763:	contig of 4175 bp in length	/note="assembly_fragment:00201
35764	35863:	gap of 100 bp	fragment_chain:1"
35864	37963:	contig of 2100 bp in length	12788. .15386
37964	38063:	gap of 100 bp	/note="assembly_fragment:00309
38064	44724:	contig of 6661 bp in length	fragment_chain:1"
44725	44824:	gap of 100 bp	15487. .21116
44825	49012:	contig of 4188 bp in length	/note="assembly_fragment:00082
49013	49112:	gap of 100 bp	fragment_chain:2"
49113	56246:	contig of 7134 bp in length	21217. .27096
56247	56346:	gap of 100 bp	/note="assembly_fragment:00610
56347	58741:	contig of 2395 bp in length	fragment_chain:2"
58742	58841:	gap of 100 bp	27197. .31488
58842	62628:	contig of 3787 bp in length	/note="assembly_fragment:00367
62629	62728:	gap of 100 bp	fragment_chain:2"
62729	73857:	contig of 11129 bp in length	31589. .35763
73858	73957:	gap of 100 bp	/note="assembly_fragment:00292
73958	76876:	contig of 2919 bp in length	fragment_chain:3"
76877	76976:	gap of 100 bp	35864. .37963
76977	83827:	contig of 6851 bp in length	/note="assembly_fragment:00675
83828	83927:	gap of 100 bp	fragment_chain:3"
83928	92033:	contig of 8106 bp in length	38064. .44724
92034	92133:	gap of 100 bp	/note="assembly_fragment:00404
92134	95917:	contig of 3784 bp in length	fragment_chain:5"
95918	96017:	gap of 100 bp	44825. .49012
96018	104147:	contig of 8130 bp in length	/note="assembly_fragment:00317
104148	104247:	gap of 100 bp	49113. .56246
104248	108202:	contig of 3955 bp in length	fragment_chain:4"
108203	108302:	gap of 100 bp	/note="assembly_fragment:00033
108303	130167:	contig of 21865 bp in length	fragment_chain:4"
130168	130267:	gap of 100 bp	56347. .58741
130268	132754:	contig of 2487 bp in length	/note="assembly_fragment:00302
132755	132854:	gap of 100 bp	fragment_chain:4"
132855	136967:	contig of 4113 bp in length	58842. .62628
136968	137067:	gap of 100 bp	/note="assembly_fragment:00369
137068	146448:	contig of 9381 bp in length	fragment_chain:5"
146449	146548:	gap of 100 bp	62729. .73857
146549	150886:	contig of 4338 bp in length	/note="assembly_fragment:00424
150887	150986:	gap of 100 bp	73958. .76876
150987	153567:	contig of 2581 bp in length	fragment_chain:5"
153568	153667:	gap of 100 bp	/note="assembly_fragment:00469
153668	155805:	contig of 2138 bp in length	76977. .83827
155806	155905:	gap of 100 bp	fragment_chain:5"
155906	158786:	contig of 2881 bp in length	/note="assembly_fragment:00581
158787	158886:	gap of 100 bp	fragment_chain:6"
158887	166773:	contig of 7887 bp in length	83928. .92033
166774	166873:	gap of 100 bp	/note="assembly_fragment:00518
166874	175651:	contig of 8778 bp in length	fragment_chain:6"
175652	175751:	gap of 100 bp	92134. .95917
175752	179746:	contig of 3995 bp in length	/note="assembly_fragment:00059
179747	179846:	gap of 100 bp	fragment_chain:6"
179847	182392:	contig of 2546 bp in length	96018. .104147
182393	182492:	gap of 100 bp	/note="assembly_fragment:00137
182493	186389:	contig of 3897 bp in length	fragment_chain:7"
186390	186489:	gap of 100 bp	104248. .108202
186490	188775:	contig of 2286 bp in length	/note="assembly_fragment:00510
188776	188875:	gap of 100 bp	fragment_chain:7"
188876	194459:	contig of 5594 bp in length	108303. .130167
194459	194569:	gap of 100 bp	/note="assembly_fragment:00213
194570	199301:	contig of 4732 bp in length	fragment_chain:8"
199301	199301:	contig of 4732 bp in length	130268. .132754
199301	199301:	contig of 4732 bp in length	/note="assembly_fragment:00313

